DOCUMENT RESUME

ED 214 873 SP 019 810

AUTHOR Sirotnik, Kenneth A.
TITLE Psychometric Implications of the Unit-of-Analysis "Problem" (With Examples from the Measurement of Organizational Climate). A Study of Schooling in the United States. Technical Report Series, No. 3.
INSTITUTION California Univ., Los Angeles. Graduate School of Education
SPONS AGENCY Institute for Development of Educational Activities, Dayton, Ohio.
PUB DATE 79
NOTE 79p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Classroom Environment; Correlation; Educational Research; *Interaction; *Measurement Objectives; Measures (Individuals); *Psychometrics; Research Design; *Research Methodology; *Statistical Analysis; Test Interpretation; Test Theory
IDENTIFIERS Study of Schooling (A)

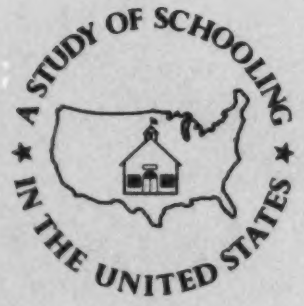
ABSTRACT

The thesis of this paper is that the decision to use one of three approaches to unit-of-analysis in educational research should be based on substantive considerations, not statistical factors. In addition to the commonly used "total analysis" (regression analysis across individuals), the within and between analyses are inherent in the same data. The within analysis is defined as an individual-based approach in which group differences are held constant, computed by correlation across individuals within each group separately, and averaging or "pooling" correlations if appropriate. The between approach uses the group as the unit-of-analysis by first computing means on the variables across individuals in each group, then using these group means to compute correlations across groups. To illustrate the uses of within and between approaches, a data-based discussion of the conceptual differences is provided in five sections. The first section consists of a review of unit-of-analysis literature, and a presentation of between, within, and total results for the correlations between two variables is included. The second section examines statistical properties of the three approaches when doing internal consistency analyses for constructs assumed to be unidimensional or when exploring the multidimensionality of a set of items using factor or cluster analysis techniques. In the third section, a literature review on methods of measuring organizational climate is presented. The fourth and fifth sections contrast the between and within approaches and discuss total, between, and within psychometric analyses for measuring climate variables using data collected in "A Study of Schooling." Appended are a list of 42 references and 18 tables presenting information about total, between, and within analyses. (FG)

ED214873

A STUDY OF SCHOOLING

Technical Report Series



Distributed by the Laboratory in
School and Community Education

Graduate School of Education
University of California
Los Angeles, California 90024

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John I. Goodlad

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SP 019 810

SUPPORTED BY

The Danforth Foundation
International Paper Company Foundation
The JDR 3rd Fund
Martha Holden Jennings Foundation
Charles F. Kettering Foundation
Lilly Endowment, Inc.
Charles Stewart Mott Foundation
National Institute of Education
The Needmor Fund
Pedamorphosis, Inc.
The Rockefeller Foundation
The Spencer Foundation
United States Office of Education

Conducted under the auspices of the Institute for
Development of Educational Activities (/I/D/E/A/)
A Division of the Charles F. Kettering Foundation

John I. Goodlad, Principal Investigator

PSYCHOMETRIC IMPLICATIONS OF THE UNIT-OF-ANALYSIS "PROBLEM" (WITH EXAMPLES FROM THE MEASUREMENT OF ORGANIZATIONAL CLIMATE)

Kenneth A. Sirotnik

Technical Report No. 3

1979

A Study of Schooling is based upon the assumption that improving schools requires knowing what is happening in and around them. A comprehensive data-base of contextual information was obtained from students, teachers, administrators, parents and observers at all grade levels in thirty-eight elementary and secondary purposively sampled schools. *It is strongly recommended that readers of any technical report in this series first read Technical Report No. 1 which outlines the details, scope and limitations of the Study as a whole.*

It must be understood that this series of technical reports does not constitute the Study. Some reports are highly specitic "molecular" inquiries while others take a more "molar" view across data sources, schooling levels, etc. Some reports are more methodological in nature arising out of issues in data analysis. Many of the reports quite naturally overlap in data analysed and interpretations rendered. Some authors have approached their task as consisting mostly of data description with little discussion beyond the presentation of the data. Others have ventured further into the realm of interpretation and speculation. *It must be further understood that data-based inferences can and do differ among researchers who come at the data from differing points-of-view.* Authors, therefore, are duly acknowledged for each report and are responsible for the material presented therein.

Psychometric Implications of the Unit-of-Analysis "Problem"
(With Examples from the Measurement of Organizational Climate)

Introduction

It will be useful for the purposes of this paper to distinguish between two analytic phases of the research process which I will call the psychometric and the study phases. The former phase refers to those analyses aimed at assessing the qualities of the measurement devices. The latter phase refers to those efforts which take the measures as given and pursue statistical analyses relevant to the research questions. Now research methodologists in education have for some time been concerned with the need to carefully identify the appropriate unit of analysis, but primarily in the study phase and in the context of experimental and quasi-experimental designs involving intact classrooms. Yet, when it comes to practice, researchers have persisted in performing regression analyses (usually in the form of simple correlations or analyses of variance) across individuals, even though these individuals are nested in one or more grouping factors directly relevant to the investigation. Furthermore, and of particular interest to the present paper, when such grouping factors exist and empirical investigation of the dimensionality in a set of items is of interest--the psychometric phase--item analyses and/or factor analyses inevitably take place using covariance and/or correlation matrices computed over all individuals regardless of group membership.

Computing indeces of relationships between variables across individuals ignoring possibly relevant grouping factors can be referred to as the total analysis. But there are two other very distinct types of analyses inherent in the same data, viz., the within and between analyses. The within analysis is also an individually based approach, but one in which group differences are held constant. This is accomplished by computing correlations across individuals within each group separately and, if appropriate, "averaging" these correlations across groups for a pooled within analysis.[2] The between approach, on the other hand, uses the group as the unit-of-analysis. This is accomplished by first computing means on the variables (other aggregate measures certainly are possible but rarely used) across individuals within each group. Then correlations are computed across groups using these group means. It turns out that the within and between approaches can produce very different results, with the total analysis being a combination of the two. (This follows directly from the well-known partitioning of the total sum of squares into statistically independent between and within additive component sums of squares.)

The unit-of-analysis "problem", then, becomes one of deciding which one (or more) of these three basic approaches to use. The arguments to be advanced here attempt to support the following conclusions: The unit-of-analysis problem is not a statistical problem, and decisions must be based upon substantive considerations. For psychometric applications, these considerations ultimately end up focusing on the operational role of the item, i.e., what property of what object it is really intended to measure. In essence, if the property is viewed as fundamentally systemic,

then the between analysis is most appropriate. The averaging process across individuals within groups is nothing more than an operational device for arriving at a single systemic measurement. If the property is viewed as fundamentally phenomenological, given a group or systemic context, then the within analysis is most appropriate. In this case, although the group mean can still be a relevant statistic, it must be conceptualized as an average of individual perceptions with its meaning being derived from the meanings of those individual perceptions. The total analysis becomes relevant only when grouping factors become irrelevant.

Having exposed these conclusions at the outset, I hope they do not produce the reaction, "Of course...so what's all the fuss about?" For in many ways, these conclusions are self-evident. Yet for some reasons (I will speculate on some later on) they have been largely ignored by researchers. Apparently, there is more to it than meets the eye. I have personally found it easier to talk about unit-of-analysis issues than to contend with them in the course of conducting scale analyses for certain instruments being used in the Study of Schooling at IDEA. For these reasons, I have resisted the temptation to write a shorter and more abstract paper on the conceptual differences between between and within approaches. It would be a disservice to the reader and to the basic theme of this paper not to thoroughly illustrate the problem with some research area and with the results of a data-based attempt to look at the potential differences in approach. I have therefore organized the remaining paper into the following sections:

1. An abbreviated review of unit-of-analysis literature with a simple presentation of total, between and within results for the correlation between two variables.

2. A not very technical look at the statistical properties of total, between and within approaches when doing internal consistency analyses for constructs assumed to be uni-dimensional or when exploring the multi-dimensionality of a set of items using factor and/or cluster analysis techniques.
3. A review of the literature on measuring "organizational climate" with the express purpose of looking for the connection between what is being operationalized and how psychometric analyses are designed.
4. A general discussion exploring conceptual differences while contrasting between and within approaches, using the measurement of organizational climate as a case in point, but hopefully with implications for any substantive area.
5. A presentation and discussion of total, between, and within psychometric analyses for "climate-like" items using data collected in the Study of Schooling.

Selected Literature Review and Basic Concepts for Bivariate Relationships

Sociologists, when interested in correlations between variables at the individual level, often have data available only at aggregated levels (e.g., communities, social entities, etc.). The practice of using correlations computed over aggregates ("ecologies") as estimates of correlations at the individual level was substantially discouraged by Robinson (1950) who mathematically demonstrated the anomalous results that can occur when correlating within groups, between groups, and totally across individuals irrespective of groups. In essence, Robinson derived the functional relationship hooking together the total, within and between correlations and the magnitude of mean differences between the groups.[3] Following Knapp's (1977)

notation, this functional relationship in terms of population parameters can be written as follows:

$$\rho_t = \eta_X \eta_Y \rho_b + \sqrt{1-\eta_X^2}\sqrt{1-\eta_Y^2}\,\rho_w$$

Suppose a population of N individuals having scores on variables X and Y is aggregated into g groups with n_j (j = 1,g) individuals in each group. If the means on variables X and Y are computed for each group, μ_{Xj} and μ_{Yj} respectively, then deviation scores on X and Y can be computed for each individual relative to his/her group mean, i.e., for individual i in group j, $x_{ij} = X_{ij} - \mu_{Xj}$ and $y_{ij} = Y_{ij} - \mu_{Yj}$. Now the parameters in the above equation can be defined as follows:

ρ_t = the <u>total</u> correlation computed across the N pairs (X_{ij}, Y_{ij}) of scores.

ρ_b = the <u>between</u> correlation computed across the g weighted (by n_j) pairs (μ_{Xj}, μ_{Yj}) of group means.[4]

ρ_w = the <u>within</u> correlation computed across the N pairs (x_{ij}, y_{ij}) of deviation scores.

η_X^2 = the correlation ratio (eta) between groups for variable X, i.e., the ratio of between-group variation to total variation measuring the relative magnitude of mean differences between groups.[4]

η_Y^2 = the correlation ratio between groups for variable Y.

In the case of minimal between group differences (i.e., $\eta_X^2 \doteq \eta_Y^2 \doteq 0$), the value of the total correlation (ρ_t) approaches that of the within correlation (ρ_w). In the case of minimal within group differences (i.e., $\eta_X^2 \doteq \eta_Y^2 \doteq 1$) the value of the total correlation approaches that of the between correlation (ρ_b). In the more interesting and more typical cases of moderate values of between and within variation, all sorts of possibilities occur.

The data in Tables 1a-d, although simplistic and contrived,

- - - - - - - - - - - - - -
Insert Tables 1a-d here
- - - - - - - - - - - - - -

sample a variety of within-between-total correlational outcomes and should give pause to any researcher who has arbitrarily selected units for analyses during the psychometric or study phases. Holding the correlation ratios relatively constant, these tables demonstrate quite different between and within correlation outcomes, each with their predictable effect on the total correlation. Tables 1a and 1b show how substantial between group correlation can be obtained under both conditions of extreme homogeneity and heterogeneity of within group correlations. Tables 1c and 1d demonstrate the same result for zero between group correlation.

Still concerned mainly with the problem of estimating individual correlations from aggregated data, rigorous discussions of the unit-of-analysis problem entered the educational literature in the mid 70's, e.g., Burstein (1975) and Sockloff (1975). But often in educational research studies, data are, in fact, available at two or more potential levels of analyses such as classrooms, schools, districts, etc. And although the mathematical formulations of the unit-of-analysis problem remain the same, the "problem" becomes refocused--on selecting appropriate levels for analysis rather than on estimating from one level of analysis to another.

Over a decade ago, the substantive implications of data availability at different analytical levels was recognized in the psychological literature. Norman (1967), anticipating present-day methods, partitioned the covariation in fairly complex analysis of variance designs and showed how potentially different correlation coefficients existed for each source of variation as well as for pooled sources of variation. Using Norman's (1967) expanded design, Marks (1970)

generalized Robinson's (1950) bivariate formulation to the multivariate case. The lesson to be learned from these inquiries was aptly stated by Norman:

> "How highly related are variables X and Y?" is a fundamental and ubiquitous question in empirical science. It is also, despite its seeming simplicity and directness, a highly complicated one. A major tenet of the following argument is, in fact, that this apparently singular and straightforward query, as it arises in typical research settings, actually entails an entire set or multiplicity of separate and logically independent forms of the question. It is further argued that the kinds of data ordinarily collected in even the most rudimentary empirical studies of such relationships provide a basis for computing separate and appropriate answers to each of the possible variants of the question, and that, in general, these several answers will not be the same. (p. 273)

More currently, this general theme has been considerably elaborated upon by Cronbach (1976). Although the substance of Cronbach's arguments was addressed to the study of aptitude-by-treatment interaction, the implications are general:

> If any fraction of the argument herein is correct, educational research--and a great deal of social science--is in serious trouble. The implication of my analysis can be put bluntly... The majority of studies of educational effects--whether classroom experiments, or evaluations of programs, or surveys--have collected and analyzed data in ways that conceal more than they reveal. The established methods have generated false conclusions in many studies. (p. 1)

In essence, Cronbach convincingly argues that correlations between aptitude and treatment outcomes using class means only do not tell the whole story--within class correlations must be considered simultaneously and, most importantly, correlations computed across individuals across classes are hopeless indices of confounded between and within effects.

Perhaps the most important lesson to be learned is that the so-called unit-of-analysis "problem" is more apparent than real, viz., that it is not a unit-of-<u>analysis</u> problem at all, rather it is a "unit-of-<u>selection</u>" problem, the resolution of which is wholly substantive and contingent upon the concepts and the purposes behind the research questions. This means that research questions to be investigated during the study phase cannot be merely clear grammatical statements about the relationship between

two or more constructs. The researcher must also be able to reasonably answer such questions as "Why am I asking this question?" and "What might I do with the results of my inquiry?" Once the researcher's points-of-view and purposes are sufficiently understood, the unit(s) (more than one can certainly be relevant) are selected and then data are analysed in terms of these units.[5] The implications of this kind of understanding for analysis during the psychometric phase translates directly to understanding the purpose of measurement at the item level. Moreover, whether or not psychometric analysis rests on either between or within approaches, group means could still be used during the study phase. To resolve the unit-of-analysis issue at the psychometric phase, therefore, the researcher must get right down to the nitty gritty of what each item is intended to operationalize.

I have seen no discussions in the literature attempting to reconcile unit-of-analysis issues at the psychometric and study phases of data analysis, except for some beginning inquiries by Cronbach (1976):

> Once the question of units is raised, all empirical test construction and item-analysis procedures need to be reconsidered. Is it better to retain items that correlate across classes? or items that correlate within classes? A correlation based on deviation scores within classes indicates whether students who comprehended one point better than most students also comprehended the second point better than most--instruction being held constant. A correlation between classes indicates whether a class that learned one thing learned another, but this depends first and foremost on what teachers assigned and emphasized. It is the items that teachers give different weight to that have the greatest variance across classes. If some teachers work hard to teach use of the semicolon and others consider it unimportant, the semicolon items will correlate comparatively high over classes. If teachers who care about semicolons may or may not care about colons, a low across-classes correlation between semicolon and colon items is to be expected. This leads us to regard the between-group and within group correlations of items as conveying different information, and makes the overall correlation for classes pooled an uninterpretable blend. (pp. 9.19-9.20)

Implications of this quote and the foregoing discussion will be explored in the next two sections, first somewhat abstractly in terms of the statistical operations during psychometric analyses and then more concretely in terms of common practice in a substantive body of literature.

Total, Within, and Between Approaches in Typical Psychometric Analyses

A beginning point is with the simple test theory model which assumes a random sample of i items drawn from an infinitely sized domain of items defining a single construct of interest and a random sample of n persons drawn from an infinitely sized population of interest. This results in a completely crossed and random n-by-i analysis of variance model. Allowing the use of I and P to represent both names for item and person sources of variance as well as symbols for the associated parameter effects in the underlying linear model, Table 2 presents the components of variation estimated

- - - - - - - - - - - - - - -
Insert Table 2 about here
- - - - - - - - - - - - - - -

by the mean squares for this design. Also presented in this table is a common measure of internal consistency, coefficient alpha (Cronbach, 1951), which is of classical test theory form, i.e., a true score to observed score (true plus error scores) variance ratio. (The variance component associated with persons' scores represents true score variance and one ith of that associated with the person-by-item interaction represents the error score variance.)

But suppose the n persons are really composed of p persons organized in g groups, i.e., n = gp. In this case, we really have a 3-factor design with persons nested within groups and both persons and groups crossed with items. Table 3 presents the variance components associated

- - - - - - - - - - - - - - -
Insert Table 3 about here
- - - - - - - - - - - - - - -

with this design.[6] Clearly, two kinds of alpha coefficients can be computed. The first looks exactly like that in Table 1 in terms of its parameter definition. But its estimation is quite different since the mean squares for persons and persons-by-items are now averages across groups--this is equivalent to utilizing the pooled, within groups, covariance matrix among items in the computation of an alpha estimate, hence the notation α_{within} (or α_w) for this coefficient.

The second coefficient simply treats groups as units (analogous to persons above) having item scores equal to the mean item scores of the persons within the groups. The computation of this estimate is equivalent to utilizing the covariance matrix of item means--a between groups partitioning of the variance--and hence the designation of $\alpha_{between}$ (or α_b) for this coefficient.

Now suppose the appropriate design is given by Table 3 but the reseacher proceeds as if there existed no groups, i.e., as if the design represented by Table 2 was appropriate. In this case, Table 2 represents what this reseacher sees, but Table 4 represents what this reseacher gets. In essence, the persons and groups effects become confounded as do their

- - - - - - - - - - - - - - -
Insert Table 4 about here
- - - - - - - - - - - - - - -

corresponding item interaction effects. (The E [MS] of Table 4 are easily obtained from those in Table 3 by first converting to E [SS], pooling these for confounded variance sources, and converting back to E [MS] by dividing by the pooled degrees of freedom.)

Table 4 also shows what the researcher is getting when (s)he computes alpha using "Table 2 mentality." Clearly, the resultant ratio is of classical form i.e., true score to observed score variance. However, the

true score involved is an additive function of both person's true scores and group's true scores. Likewise, the error score is really an additive function of the person's and group's error scores.

So what does all this mean for selecting an appropriate unit of analysis? Nothing--that is, until we find out what the items are supposed to be measuring. The recurrent theme of this paper will be that all levels of analysis may be interesting, yet potentially conflicting in results--not paradoxically, but only because of implied differences in purpose and point-of-view (or conceptualization or theory) behind what might appear as the same constructs and research questions. Suppose, for example, the researcher was measuring a construct presumed to be purely individual in nature and not influenced by contexual (systemic) variables. We would therefore expect near zero variance components associated with groups (σ^2_G & σ^2_{GI}) in which case the alpha in Table 4 reduces to that associated with a simple person-by-item design (i.e., the α of Table 2).

On the other hand, suppose the researcher was measuring a construct presumed to be purely systemic in nature, the responses by persons within the group being solely a function of contexual constants. Now we would expect near zero variance components associated with persons (σ^2_P and σ^2_{PI}) in which case the alpha in Table 4 reduces to that associated with a simple group-by-item mean design (i.e., the α_b of Table 3).

But suppose the researcher presumes the items to be measuring both an individual and systemic construct. Now we would expect all non-zero variance components and we would have values for both α_w and α_b in Table 3. Furthermore, these two coefficients are independent; that is, data could easily be contrived where one coefficient was zero and the other was one (or any other combination in the 0-1 range) in much the same manner as

Table 1 was constructed. This is, of course, as it should be since each coefficient is telling us something different.

The trouble is that many of us fail to adequately sort out what it is that we want to measure for which unit(s). Countless researchers have utilized alphas conceived of in the form of Table 2 when either one or both of those in Table 3 were the relevant indices. As a result, coefficients given by Table 4 have been reported, representing a mixture of between and within group item covariation with questionable interpretive value.

Although I won't attempt a mathematical demonstration, much of what has been said is generalizable to the multivariate case. In particular, factor or cluster analyses of correlation matrices computed using either between or within approaches can yield extremely different results. Analysis of the total correlation matrix would result in a compromised structure equal to a kind of weighted average of the pooled within and between structures. (The weightings would probably be proportional to multivariate ratios of between-to-total group variance analogous to the bivariate correlation ratio eta squared.) In any case, the researcher is left once again with an apparent dilemma. Several psychometric treatments of a single set of data can potentially (and often do) yield outcomes which can have both interpretable differences and similarities. Left at the level of statistical treatment, the result appears paradoxical. But once the researcher understands that the very same data can be used to operationalize different constructs, the paradox is resolved and (s)he has only to figure out what properties of what objects (s)he is really trying to measure.

Moreover, what unit (and statistical aggregate) the researcher is planning to use during the study phase does not necessarily dictate the appropriate level of analysis during the psychometric phase. The

possibilities are sorted out in Table 5 given choices of statistical aggregates

- - - - - - - - - - - - - - -
Insert Table 5 about here
- - - - - - - - - - - - - - -

during each analytical phase. For example, the first row relates to individually-based instruments like personality inventories which might be presumed to be immune to group differences, at least for the purposes of certain investigations. If, in the case of the third row, one made the case for item means (or variances, etc.) as the sine qua non measures of the group, then it would follow that the multivariate structure among items be explored using these measures and that the group be given composite scores for emerging factors based on these measures (e.g., the average or weighted average of item means, variances, etc.). Suppose one wished to look at relationships across groups such that aggregates at the group level have substantive meaning insofar as the data (individuals' scores) comprising the aggregates have meaning relative to the given group. This leads directly to the cell determined by the intersection of the second row of the psychometric phase and third column of the study phase in Table 5. Note that Table 5 merely enumerates possibilities--it cannot guide the researcher in choosing the appropriate cell(s).

An Example: The Measurement of Organizational Climate

Although I might be able to get away with keeping this discussion at its current level of abstraction, I think it would be a disservice to the reader (and to the theme of this paper) not to test the relevance of these issues with some substantive topic of fairly wide interest. The measurement of organizational climate is well suited for this purpose, given the abundance of literature in the field and the natural nesting of

persons within groups (organizations). Moreover, it is out of this context that my concern regarding unit-of-analysis issues reached its present level of intensity.

Among the many substantive interests of the Study of Schooling currently underway at I.D.E.A. are the attitudes of teachers and students towards their respective work places.[7] Several "climate" inventories were constructed using Likert-type items to assess various dimensions of the classroom learning environment as perceived by students and of the school work environment as perceived by teachers. In the course of designing internal consistency analyses for these instruments, the several levels at which the analyses could be conducted became quite apparent. For example, data were available from the faculties of 38 schools on items pertaining to organizational climate constructs such as "trust," "role definition," "morale," "autonomy," etc. To assess the internal consistency of these a priori constructs or to explore ex post facto the dimensionality of the entire set of items, we could compute the correlation (or covariance) matrices among items

a) within each school separately and/or pool (average) these matrices
b) across schools using item means computed for the teachers at each school and/or
c) across all teachers ignoring the school factor.

After an initial period of frustration centering mostly on statistical issues contrasting the several approaches and finding the "right" analysis, we realized and accepted the co-existence of the several analyses, each requiring its own conceptual (and <u>not</u> statistical) justification.

At this point we returned to the literature on organizational climate looking specifically for how unit-of-analysis considerations during this psychometric phase were handled.[8] What we found was quite remarkable. Most researchers either implicitly or explicitly acknowledged the unit-of-analysis issue when using already developed climate scales in studies designed to test hypotheses regarding the correlates of these measures. The group was often selected as the unit when the constructs were viewed as "organizational" in nature and the individual within a single organization when the constructs were viewed as "psychological" in nature.

Ironically, however, concerns pertaining to appropriate analytic units during the instrument development phase were virtually non-existent. With only one exception (to be discussed below) none of the nearly forty reports and articles reviewed, which included discussions of psychometric analyses, included an explicit rationale for coordinating the units selected for psychometric analyses with those to be selected for subsequent analyses pertaining to research questions.

The most popular method of instrument development (used in approximately 75% of the reports reviewed) was to compile a set of items operationalizing the researcher's conceptual framework for climate. Data were then gathered from members of one or more organizations and the item intercorrelation matrix computed across individuals was factor analysed. When only one organization was involved, this becomes by default, a within group analysis; and about 40% of the factor analytic studies were of this type. When more than one organization was involved, this becomes a total (across individuals irrespective of organization membership) analysis, and the remaining 60% of the factor analytic studies were of this type.

In the remaining 25% of the reports reviewed, climate constructs were defined by items a priori, followed by typical item analysis procedures designed to increase the internal consistency of each scale. Again, the variance-covariance matrices implicit in these analyses were computed across individuals in a single organization (about one third of the studies) or across individuals across two or more organizations (about two thirds of the studies).

We were able to locate only three studies in which the dimensionality in a set of items was explored using something other than the correlation matrix computed across individuals. Taylor and Bowers (1972) scored work units (sub-groups) within an organization on each item using the subgroup's mean, correlated items across subgroups both within and across organizations, and cluster analysed the resultant matrices. However, no discussion was included as to why this approach was selected. The only study which we located which explicitly acknowledged a unit-of-analysis problem (and rationale for resolving the problem) is that contained in the manual for the College and University Environment Scales (Pace, 1963). Because central tendency was viewed as being the most relevant measure of organization climate variables, scale means were intercorrelated and factor analysed over institutions. Further item selection procedures were based on correlating item means with the resultant factors, again computed over institutions. Finally, we found one study which set out to determine differences in the dimensionality of "environmental press" in high schools as perceived by students. Using the High School Characteristics Index (Stern, 1970), Mitchell (1968) intercorrelated the 30 scale scores across all students and schools (total analysis), across students in each school separately (within analyses), and across all schools using scale means computed for each school (between

analysis). Mitchell's enlightened treatment of the results is well illustrated in this description of two of the four factors he found:

> The results of this analysis and the separate analyses described in the paragraph above seemed to suggest that the between-school and within-school factorial patterns were quite similar in nature, but that the School Activities was more reflective of between-schools variance, and hence more critical in differentiating among schools, while that associated with the Negative Attitude toward the Environment was relatively more sensitive to within-school differences most probably associated with personality differences in the respondents. (pp. 520-521.)

In sum, two studies were found which factor analysed a between correlation matrix, one (Taylor and Bowers, 1972) at the item level and one (Pace, 1963) at the scale level. A third study actually factor analysed and compared all three kinds of correlation matrices at the scale level. No studies were found which factor analysed the pooled within correlation matrix. (It is interesting to note that two studies were found, Stern, 1970 and Moos, 1972, which computed alpha coefficients for scales using pooled within item variances; and yet each of these studies also explored the dimensionality of items and scales by correlating across individuals across organizations.) Table 6 highlights the several approaches to instrument development revealed in the above literature review for the more popular climate inventories found in the industrial organization literature and for most of the climate inventories designed for teachers in schools which turned up in our literature search.

Toward a Conceptual Understanding of Between Versus Within

It seems fairly safe to draw the conclusion that the psychometric implications of unit-of-analysis issues have been almost universally ignored in the organizational climate literature. And I would guess the same conclusion might be drawn for most other substantive areas which

produce data amenable to between and within analyses and interpretations. Why? I would like to engage in two speculations, the first briefly and the second a bit more in depth.

First, acknowledging the relevance of the phenomenology of the individual perceivers within an organization and thus pursuing the analysis of the pooled within covariance structure among items measuring a certain construct opens up a "Pandora's Box" of sticky conceptual problems. Clearly the pooled within structure is but an average of separate within structures and thus representative of within structures only to the extent thay they are all relatively similar. But what if they are not all similar? What if the alpha coefficient is high for one organization but low for another? Is it only unreliability or is the construct simply not valid for that particular organization? The dilemma gets worse when exploring the dimensionality of these structures utilizing cluster or factor analysis techniques. In doing so, the researcher confronts the possible destruction of a primary working assumption, viz., the existence of a consistent set of constructs across organizations of the same generic type (e.g., schools, hospitals, insurance companies, etc.).[9] Cronbach (1976) has stated the case as follows:

> ....variables that have the same operational definition may have different substantive interpretations at the individual and the aggregate levels.... The person using factor analysis as a tool in theory construction, then, will need one set of factors for his between-groups theory and another set of factors for his within-groups theory. To be sure, he may find that the two sets of constructs coincide, but that is a possibility to be evaluated, not assumed. (p. 9.8)

I am simply suggesting that these kinds of substantive consequences are alien to most researchers and therefore consciously or subconciously avoided.

At minimum, this first speculation suggests the importance of thoroughly understanding what it is you are trying to measure and carefully operationalizing constructs in accordance with this understanding. There is a widely acknowledged but not so widely heeded aphorism that your results are never any better than the data upon which they're based. Quite often researchers spend disproportionately less time in the operationalization and instrument development phase than at any other phase in the research process. If anything, perhaps it should be the reverse. Now the informed selection of appropriate levels of analysis during psychometric studies rests primarily upon understanding what it is that is being operationalized at the item level--a process in direct conflict with the above stated belief about current research practice. This, then forms the basis of my second speculation as to why the psychometric implications of unit-of-analysis issues have been largely ignored.

This second speculation deserves exemplification and the study of organizational climate again serves this purpose rather well. Synthesizing the literature in this field proved to be an exceedingly difficult task, mainly because of the conflicting views of what, in fact, organizational climate is. Much of this confusion is manifested in debates over whether or not climate is "psychological vs. organizational," "descriptive vs. affective," and/or "job related vs. organizationally related." It turns out that these issues are not at all mutually exclusive and, most importantly, appear to be very much related to the question of which unit(s) are to be selected for analysis. (Compare, for example, the reviews by Howe and Gavin, 1974 and Jones and James, 1976). Essentially, the first issue revolves around the extent to which interpretations of variable relationships

are individually or organizationally based; and as such, some researchers have treated the distinction isomorphically with the choice of individual or group as the unit of analysis. At the core of the second issue is the degree of sentiment elicited, descriptive items at the lower end and affective items at the upper end of the continuum. Responses to descriptive items about any organization should be relatively less variable than responses to affective items. Some researchers have even suggested that climate exists only when consensus exists on descriptive variables demanding, therefore, the group as the unit of analysis. (See, for example, Guion, 1973). Perhaps the most debated issue centers on the relatively high correlations between job related and organizationally related variables and, therefore, whether or not they are really measuring one or the other or both. Payne, et. al. (1976), explicitly acknowledging the unit questions, attempted to sort out these issues. Using three dichotomous facets (individual vs. group as unit; job vs. organization related variables; and affective vs. descriptive variables), eight types emerge with job satisfaction and organizational climate being only two.

Certainly the substance of the above issues is of theoretical importance in the study of organizational climate. And so are the implications for constructing measurement devices--a seemingly critical discussion which was rarely found in the literature. For example, reviews like those of Howe and Gavin (1974) and James and Jones (1976), and synthesizing attempts like that of Payne, et. al. (1976), omitted operationalized distinctions of psychological vs. organizational, descriptive vs. affective, and satisfaction vs. climate in terms of exemplary items. Apparently a great deal of time has been spent conceptualizing and testing models of organizational climate. But how much time has been spent bridging that gap

between constructs and the empirical world? Again, it is my contention that the time spent in this latter effort is directly proportional to the ease with which dilemmas related to "between" versus "within" analysis are appropriately resolved.

Other than what is contained in the following quotation, I have seen no attempt to give serious thought to coordinating unit-of-analysis considerations at the study phase with those at the psychometric phase. Although referring to classroom climate measures at the scale level, Cronbach's (1976) analysis could easily be translated to the item level of any organizational climate measure and it is worth quoting at length at this point:

> The case for considering separately the between and within factor analyses is especially impressive when we move out of the ability domain. An example is the Learning Environment Inventory (LEI) developed by G.J. Anderson and Walberg (1974) within Harvard Project Physics. This is a collection of items describing instructional procedure and student attitudes; the student responds so as to report how he perceives his class. Items have been assembled into scales on the basis of their intercorrelations, and the scales have been factor-analyzed. Insofar as correlations arise simply from semantic overlap of items, one would expect similar joint-distribution shapes within and between classes. But the correlational structure, insofar as it reflects psychological differences, may be quite different. Within the class, the correlation reflects the correspondence of perceptions. Are the students most prone to describe the class as "apathetic" also the ones prone to describe it as "difficult"? Across classes the correlation speaks of a different phenomenon: When the students collectively describe the class as "difficult", do they also describe it as "apathetic"? The former refers to the phenomenology of the student, compared to other students rating the same events. The latter refers to behavioral differences between classes (though some of those differences are perceptual rather than objective).
>
> The purpose of the LEI is to identify differences among classrooms. For it, then, studies of scale homogeneity or scale intercorrelation should be carried out with the classroom group as unit of analysis. Studying individuals as perceivers within classrooms could be interesting, but is a problem quite separate from the measurement of environments. (pp. 9.18-9.19)

A critical part of Cronbach's argument is contained in the last paragraph. The first sentence of this paragraph identifies a purpose for analysis during the study phase. The last sentence begins to tackle the issue of what the items are intended to measure. The first sentence alone is not sufficient to identify the relevant unit(s) of analysis during the psychometric phase. Maximizing differences between groups is not necessarily what guides the selection of units for psychometric analyses--what you want to measure is the guide. Certainly a between analysis by its very nature tends to produce constructs which are sensitive to between group differences. This is as it should be when a between approach is appropriate. The within analysis by its very nature produces constructs sensitive to person differences within the same group. This is as it should be when a within approach is appropriate. Moreover, if these within constructs show variation across groups, on the average, this may be a relevant finding during the study phase. (And so might be the finding that the groups do not show this kind of variation.)

So the issue boils down to whether or not we are looking at differences between groups (classes in Cronbach's example) based on single measurements of a systemic property, or based on average, within group, person perceptions relating to a (perhaps the same) property. If it is the former, then the psychometric qualities are best looked at using a between analysis; if the latter, then within group analyses are most relevant. And in either case, looking at mean differences between groups during the study phase can be a relevant analysis given the purposes of the research. Now I will admit again that this is easier for me to say than do, particularly when it comes to operationalizing systemic constructs using the types of items usually

found in climate questionnaires. Thus, in the remainder of this section, I wish to pursue the apparent differences in measurement implied by within versus between analysis.

Consider the following items designed to measure the construct of "trust" and suppose they are Likert-type items responded to by teachers in school:

1. I am generally a trusting type of person.
2. I trust the staff members at this school.
3. We trust one another at this school.
4. You trust one another at this school.
5. Staff members trust one another at this school.

What object is being measured by each of these items? Can there be more than one? Certainly teachers are responding to each item, but are these responses measuring something about the teacher, something about the school, or both? Suppose item <u>means</u> are computed. Do these means represent measures of an intrinsic property of the school, with averaging over teachers being incidental and merely a convenient operational device for getting at this property? Or do these means represent indices of central tendency of distributions of measures of a property of teachers? Can these means represent both kinds of measures?

At first glance, the differences between these queries may appear subtle and perhaps even unimportant. But I think their resolution <u>is</u> the unit-of-analysis "problem". Understanding what you are measuring at the item level <u>is</u> selecting the appropriate unit to be analysed. It is easy to relate to items as measuring something about individuals since the whole idea and use of measurement in the behavioral sciences has come about from this perspective. It is not so easy to relate to items as measures of something

systemic. As another example, let's use "teacher influence" as a single construct for which we have a set of items and thus a scale score for each teacher in any school. Understanding the multitude of variables (e.g., sex, beliefs, length of employment, internal-external personality type, etc.) and their interactions which account for the variance in "teacher influence" scores is clearly an interesting and worthy task. So also is understanding why some schools seem to have more potential for teacher influence than others. This implies a conceptual quality or attribute about the school--or rather the "ambience" or "climate" of the school--called "teacher influence". This attribute is obviously intangible and only there because humans are there in some kind of interactive process....which is precisely why we have to ask the humans about it to get at it. (We wouldn't even think of this approach to data gathering if we were interested in something like the average daily absenteeism rate of teachers). So if we are conceptually committed to the existence of this intangible attribute and we have only our variable (for good reasons) human data, what is our best estimate of the magnitude of "teacher influence" <u>in the school</u>? The mean, being an overall level effect and cancelling out the various and sundry main effects and interaction effects of other variables, would seem to be the best choice for this purpose. It is extremely important, however, that the nature of this measure be understood as an estimate of the magnitude of an attribute of the school--<u>not of an attribute of the teachers in the school</u>. The measure, as we are now conceiving it, does not exist in different quantities for different people; rather, it exists in one quantity indicating the magnitude of a single property of the group. When we choose to assess this property indirectly through the perceptions of the people in the group, we must be assuming that a

constant indicant of this property somehow exists in each person, operationalized by the overall level effect, i.e., the mean. The teachers are merely a vehicle for getting the measure, and to ascribe the school average as a measure of how any teacher perceives the school would be clearly fallacious given this line of reasoning. But can we operationalize both individual and systemic properties in the same item? Only by fiat, hopefully based upon some conceptual rationale. There is simply no logical reason (aside from substantive considerations) to suppose that the "something" being measured at the systemic level is the same thing being measured at the individual level. Philosophers of science have recognized these potential errors in reasoning for years, albeit the emphasis has been placed more on semantic ambiguities rather than fundamental differences in constructs:

> The fallacy of composition frequently occurs when we reason from the properties of elements or individuals to the properties of the wholes which they constitute. For the same word may have a different significance when applied to a totality than it has when applied to an element. Thus the fact that the soldiers of a given regiment are all "strong" does not justify the conclusion that the regiment which they constitute is "strong". The word "strong" does not mean the same in the two cases. The fact that the soldiers are Irish fails to prove that the regiment is Irish. It may be part of the British or even of the French army.
>
> The fallacy of division is the converse error of reasoning as if the properties of any whole are always properties of each part. If we say that mankind generally attains its end after trial and error, it does not follow that any individual or group will finally be successful by this method. The fact that the roman senate was a wise body does not prove that any individual in it was wise. Nor does the foolishness of assemblies, as judged by their resolutions, prove that the individual members were foolish. Men have different characters when in a group than when alone. (Cohen & Nagel, 1934, P. 377)

What, then, are some of the item characteristics that might serve to distinguish between systemic-and person-focused measurement? Consider again the four items above. The first "I"-form item is clearly non-systemic,

having no group referent and being the typical personality-type item. The remaining three items have group referents but vary in degree of potential systemic effect, with the "I"-form item most person-oriented, the "they"-form item most group-oriented, and the "we"-form item somewhere in between. In the organizational climate literature, this grammatical sequence (1st person singular to 3rd person plural) might be thought of as operationalizing the confounded continuums of psychological-to-organizational climate and affective-to-descriptive measures. In some ways, however, I think researchers in organizational climate have unwittingly camouflaged the measurement issue by creating the dichotomy between "affective" and "descriptive" items using what are essentially perceptual data. On my own "affective-descriptive" continuum, an "I-form" item like number 2 above is somewhere near the "affective" end point and a "they-form" item like number 4 is somewhere in the middle. For me, items which are nearer the "descriptive" end point might be something like these: "Most of the staff in this school are over the age of 21" or "The curriculum structure of this school is departmentalized." Obviously, such data can be gathered by observation, document review and simple demographic survey. As soon as the constructs are such that we have to question the humans in the organization, we open ourselves up to sizeable between and within variance components. I have been unable to write items reflecting constructs like "trust", "influence", "role definition", "morale", etc. which I would place near the "descriptive" end of the continuum. Such is the dilemma faced by organizational climate researchers who demand high degrees of consensus on items which are every bit as "affective" as they are "descriptive".

So the only real dilemma in selecting units to be analysed during the psychometric phase is due to the primitive state of our measurement capabilities. In the next and last section of this paper, I wish to share what we primitives at IDEA have been doing to cope with these measurement and unit-of-analysis issues.

Another Example: Measuring "Teacher Influence" in Elementary and Secondary Schools

Among the several climate-like instruments assembled for IDEA's Study of Schooling is a 33-item "teacher influence" questionnaire which is largely adopted from an instrument found to be useful in IDEA's previous Study of Educational Change and School Improvement (Bentzen, 1974). These items were written so as to sample the range of likelihood of teacher "power", from class-specific decisions (e.g., standards of pupil behavior in their own classrooms) through school-general decisions (e.g., daily schedule of whole school) to higher level administrative decisions (e.g., teacher selection and evaluation). Teachers at each school responded to each item on a 3-point "influence" scale defining "a lot", "some", and "no" influence. Table 7 lists these items in the format presented to teachers. These items are essentially third person plural items that would be defined

- - - - - - - - - - - - - - -
Insert Table 7 about here
- - - - - - - - - - - - - - -

as descriptive (vs. affective) by most climate researchers.

But for reasons already touched upon towards the end of the previous section, we have concluded that the term "climate" (not withstanding its metaphorical utility) has taken on an unfortunate connotation of environmental fact and consensus in the research literature. Risking epistemological debate, it becomes useful to distinguish fact from perception

in evaluating the kinds of data we have collected. The former term is meant simply to refer to those situational or circumstantial kinds of data resulting from the kind of responding for which we assume all respondents (real or hypothetical) are using (or would be using) the same basis for judgment. Variation in responses is assumed to be more a function of measurement error than of systematic (although not necessarily explainable) differences in interpretation. The latter term refers mostly to data of an "attitudinal" nature where variation is expected and, for reasonably reliable instruments, assumed to be due mostly to differences in individual judgment and interpretation. (Similar distinctions have been made by other climate researchers; see, for example Howe and Gavin's (1974) public vs. private standards and Walberg, Sorenson and Fischbach's (1972) ecology vs. ambience discussions.)

To exemplify this troublesome but useful dichotomy, suppose one measures the temperature of a room with a thermometer. The resultant reading (or perhaps the mean reading of several similarly placed thermometers) would fit the "factual" category. But suppose one wanted to assess how hot or cold it felt to the people in the room. One might similarly position people in the room and have them record their impressions on a hot-to-temperate-to-cold scale, yielding data which would fit the "perception-based" category. In the latter case, the distribution of impressions tells the whole story; and regardless of the central tendency, variability and shape of this distribution, it reflects something quite apart from (although not necessarily uninfluenced by) the thermometer readings.

More to the point, teacher demography, biography, number of classroom preparations, job turnover; student demography, absenteeism, dropout rate; school age, size, per pupil expenditure, departmentalization, principal turnover; district tax base; community demography.....are examples of the

"factual" variety. "Perceptual" measures are exemplified by 100% of the "climate" inventories (including our own) that have been reviewed. These typically require members of the organization to provide data on constructs such as autonomy, cooperation, motivation/reward, task emphasis, trust, role clarification, morale, conflict resolution, satisfaction, influence, etc. Whether items relating to these constructs are written in highly personalized or highly depersonalized forms, they are still affect-bound and yield data conceptually distinct from those classified in the "factual" category. The trouble starts when perception-based items are treated as having produced factual-type data on the average. It should be clear from Table 8 that even our highly depersonalized items yield substantial variance

- - - - - - - - - - - - - - -
Insert Table 8 about here
- - - - - - - - - - - - - - -

components for within school effects; moreover, between school differences (although statistically significant due to the large sample sizes) never account for more than 25% of the total variation. Therefore, we prefer to use terms like organizational "climate" or "work environment" to refer to the total picture that begins to emerge as all relevant data are analysed, both factual and perceptual varieties. The latter variety continues to be the main focus of this paper. It should be understood then, that when terms like climate or work environment are used, they are abbreviations for the phrase "selected aspects of organizational climate or work environment as measured by perception-based items responded to by members of the organization."

Now after data were collected from teachers in 38 schools, we were faced with a number of psychometric options. We could, for example, engage in

the traditional practice of checking item-scale correlations and building reasonably sized alpha coefficients for the a priori defined constructs--a not too difficult task given enough positively correlated items. Or, we could be more daring and confront the probable destruction of one or more constructs by tossing all the items into the factor analysis pot and let fall what may. In fact we opted to do both, namely, to combine both rational and empirical methods to arrive at a reasonable compromise between theory and operation relying to a large extent on the results obtained by cluster analysing the absolute values of inter-item correlation coefficients (using Hartigan's, 1975, average linkage rule to amalgamate clusters as programmed in BMDP1M, Dixon and Brown, 1977.) But which correlation matrix do we choose--the within, between or total? Given the foregoing discussion, choosing the total matrix is clearly the poorest choice from the standpoint of utility and conceptual understanding although it might be the easiest choice from the perspective of data organizaion, availability of computer programs, and conformity to common practice.

Choosing between the remaining two alternatives was not so easy, at least not for those of us at IDEA who had committed ourselves to struggling with this unit-of-analysis issue. As might be expected, two "camps" emerged. The majority of staff, comprising the "within camp", argued that our perceptual-type items are an admission of the importance we have always placed upon the phenomenology of the individual perceivers. In effect, our meaning is operationally derived from the teacher's experience within the school and how this experience interacts with the items on our questionnaire in producing teacher responses. This position fits quite nicely with another point of contention, namely the exclusive reliance upon the arithmetical

average of perceptions to stand in for what is meant by any climate construct. Given the nice mathematical properties of the mean, it is understandable how all of the technical work on unit-of-anlaysis issues has focused exclusively on this method of aggregation. However, it has also focused attention away from other, perhaps equally or more useful measures. For example, taking a very operationally oriented position, we might define any particular dimension of organizational climate as the distribution of responses obtained on the scale representing the dimension. Three classes of aggregate statistics emerge, therefore, as being potentially relevant: measures of central tendency, variability, and shape.[10] (The idealized distributions in Figure 1 provide examples of how this definition might be translated into descriptions

- - - - - - - - - - - - - - - - -
Insert Figure 1 about here
- - - - - - - - - - - - - - - - -

of organizational climate.) In any event, given this reliance upon the distribution, we are led once again to the importance being placed upon the teacher's perception within the school.

The minority position taken by the "between camp" argued that the concept of climate demanded systemic or contextual variables which are best operationally defined by the empirical associations produced by teacher consensus across schools. Certainly this camp was troubled by the sizable within school item variances. Nevertheless, sufficient between school item variances existed as well offering the potential for interpretable between school dimensions. This put into focus once again the necessity--but not the sufficiency--of substantial variance components as a determinant in the choice of appropriate level of analysis. Thus the "between camp", acknowledging their lack of factual-type data, still wished to pursue the multivariate item structure of average teacher perceptions across schools.

I must admit that I have created these two "camps" a bit artificially for expository purposes. To be sure, there are those of us in the Study of Schooling who tend to think more like clinical psychologists and naturally gravitate more towards within-school analyses; and those of us who operate more like social psychologists are drawn more towards between-school analyses. Nevertheless, we all recognize the utility of both approaches in effecting the ultimate goal of school improvement. We were thus eager to see what our data looked like empirically from both analytic perspectives.

From the national study conducted in the Spring and Fall of 1977, data are available on 664 teachers from 13 high schools, 400 teachers from 12 junior high (or middle) schools, and 286 teachers from 13 elementary schools. Actually, twelve different correlation matrices for the 33 teacher influence items were computed. The following schematic sorts out the various types of correlation matrices, samples over which they were computed, and unit of analysis:

Level of School

Correlation Matrix Type	All...	High...	Junior/Middle...	Elementary...
Total	School Teachers (N=1350)	School Teachers (N=664)	School Teachers (N=400)	School Teachers (N=286)
Between[11]	Schools (N=35)	Schools (N=13)	Schools (N=12)	Schools (N=13)
Pooled Within	School Teachers (N=1350)	School Teachers (N=664)	School Teachers (N=400)	School Teachers (N=286)

I will first contrast the between and pooled within analyses, saving until later the total analysis in order to see what would have resulted had we proceeded along traditional lines.

In order to get some initial idea of the order of dimensionality among items, the eigenvalues were plotted against number of factors for both the between and pooled within correlation matrices. (See Figure 2.) In

- - - - - - - - - - - - - - - -
Insert Figure 2 about here
- - - - - - - - - - - - - - - -

both cases it is evident that there are 3 major factors and a potential for interpreting perhaps as many as 9 factors. The range of 3 to 9 factors in the pooled within solution accounts for 33% to 53% of the total item variance. As expected (given substantially high item correlations) the between solution accounts for 20% to 30% more of the variance in this range.[12]

Tables 9 and 10 present the cluster analyses of the between and pooled

- - - - - - - - - - - - - - - - - -
Insert Tables 9 & 10 about here
- - - - - - - - - - - - - - - - - -

within correlation matrices and Tables 11 and 12 present the rotated factor

- - - - - - - - - - - - - - - - - - - -
Insert Tables 11 & 12 about here
- - - - - - - - - - - - - - - - - - - -

loadings for several orthogonal solutions for these correlation matrices.[13] A visual inspection of the cluster analyses of both correlation matrices reveals findings compatible with the analysis of their eigenvalues, viz., three major clusters with a number of potentially interpretable sub-clusters. Looking at the two cluster analyses side-by-side, both the similarities and the differences in clustering tendencies can be seen. The general trends in both analyses for "school in general" and "selection and evaluation" clusters is unmistakable; and, although less clear, the "budget" and "staff meeting" clusters and to some extent both "curriculum and instruction" and "pupil-focused" clusters are evident in both analyses.

But the hierarchical patterns evidenced by the ways in which these sub-clusters become associated with their primary clusters is quite different

in each analysis. To see this most clearly it is helpful to interpret the factor structures for one correlation matrix and then go back for a comparitive interpretation of the other correlation matrix.

Looking first at the pooled within correlation matrix, a 3-factor structure can be identified under these major headings:

A. Issues related to general conduct of school. This is a potpourri of items pertaining to community relations, extracurricular staff committments, procedures and content of staff meetings, and general student issue (e.g., dress codes and scheduling). (The cluster analysis suggest that other school-general concerns like communication with parents and staff dress codes and social gatherings could be included in this dimension).

B. Issues related to budget and the selection, assignment and evaluation of teaching and administrative staff. (The cluster analysis suggest that the other fiscal management item along with the remaining selection and evaluation items could be included in this dimension).

C. Issues related to instruction and pupil behavior which are primarily classroom focused. (The cluster analysis suggests that the one item pertaining to curriculum could also be included in this dimension).

As more and more factors are interpreted, more of the sub-clustering tendencies are uncovered. The 4-factor solution leaves factor C intact and modifies factor B slightly making it a staff selection and evaluation dimension; however, factor A is clearly separated into two dimensions: those items mostly pertaining to community-related issues (factor A) and those pertaining mostly to staff meetings (factor D). When 7 factors are rotated it is apparent that factors A, C and D have remained pretty much intact and that factor B is now a "professional staff selection and evaluation" dimension with "teaching assistant selection and evaluation" items clearly forming a separate dimension (factor F). The remaining two

dimensions are mainly defined by the two "dress standards" and the two "budget" items (factors E and G respectively). Finally, as many as 9 factors can be rotated and clearly interpreted for the pooled within correlation matrix. Factors, A, B, and D through G remain intact and factor C now consists only of those items pertaining to "pupil behavior". The remaining factors pick up the "curriculum and instruction" cluster and the "parent communication" cluster (factors H and I respectively).

In sum, based on the cluster & factor analyses of the pooled within correlation matrix, the teacher influence items can be quite clearly sorted into nine dimensions which can be organized into a hierarchical clustering pattern as follows (letters correspond to 9-factor solution):

Primarily school related issues:

- General School/community (A)
- Staff meetings (D)
- Dress standards (E)
- Parents (I)

Selection & evaluation of staff and fiscal management:

- Professional staff (B)
- Teaching assistants (F)
- Budget (G)

Primarily classroom related issues:

- Pupil behavior (C)
- Curriculum and instruction (H)

(Note that many of these item groupings correspond with our a priori categories used to generate the "influence" items.)

Following this same analytic approach to understanding the dimensionality of the between correlation matrix, several general themes become apparent: (1) despite the obvious similarities in small item clusters, there exist important changes in item composition for a number of factors, and (2) the ease with which item clusters can be analytically identified and interpreted in the pooled within analysis is substantially diminished in the between analysis.[14] Comparing both the cluster and factor analyses of the between matrix with those for the pooled within matrix, we see, at the 3-dimensional level, a merging of pupil-related and general school-related issues into one global dimension (factor A), a clear separation of curriculum, instruction and budget items into a seaparate dimension (factor C), and a relatively undisturbed staff selection and evaluation dimension (factor B).

At this point it becomes interesting to speculate on why these changes occur in moving from within to between approaches. For example, perhaps within schools teachers perceive their influence over pupil behavior much like they do for other issues which have the most direct impact upon their primary daily function, viz., classroom instruction. Averaging out these perceptions in the between analysis, we obtain a school measure of teacher influence in decisions concerning pupil behavior which may very well co-vary with discipline as a general school problem and thus co-vary with the host of other school-level measures of general school and community related issues. Similarly, issues pertaining to curriculum and instruction, which appeared to be classroom focused in the pooled within analysis, now appear to be more a matter of resources, co-varying with the degree to which staffs have more or less influence in preparing and managing the budget.

Pursuing these comparisons in the interpretation of more than three dimensions, we continue to find subtle but interesting differences. For example, global factor A becomes similar to the "school/community" factor of the pooled within analysis, the remaining items (factor C) resembling the "pupil behavior" factor in the pooled within analysis. But it really isn't this clear cut. Factor A in the between analysis retains two key items pertaining to "staff scheduling" issues (items 13 and 17), giving this dimension more of a flavor of "extra-curricular staff involvements & commitments". Factor C is not limited only to behavioral issues but consists of almost all pupil-focused items. The between dimension least similar to any within dimension is factor E. Perhaps this factor reflects the "permissiveness" of school-level policies with regard to students, with parent conferences being the inevitable result of student dress code violations. For teachers within schools, however, communication with parents is much more a function of individual pupils and class-specific issues. (Although not X'd in Table 12, item 16 loads relatively high, better than .4, on factor C in the pooled within analyses). Another difference worth noting is the relative cohesion of the selection/evaluation factor in the between analysis in contrast to the clear split into professional staff and teaching assistant components in the pooled within analysis. Given fairly pervasive but different selection and evaluation policies from district to district, we might expect all such items to be lumped together in an analysis sensitive only to those differences. However, holding school (and thus to a large extent district) differences constant (as we do in the pooled within analysis), we might expect teacher perceptions of influence in selection & evaluation to covary more along the split between their

"power" over what goes on in the classroom versus what goes on at the general school level.

The 6-factor solution for the between matrix ends up being the highest order of dimensionality interpretable using factor analysis. Combining these results with those of the cluster analysis, 6 subsets of items can be organized into a hierarchical clustering pattern as follows:

Global School/Community/Pupil Related issues:

Extra-curricular staff involvement & commitment (A)

Pupils (C)

Procedure/content of staff meetings (D)

Standards (pupil dress) and parent conferences (E)

Selection and evaluation of staff(B)

Curriculum, instruction and budget (F)

(Note that these item groupings do not correspond as closely to the a priori categories as did those for the pooled within analysis.)

Certainly the foregoing interpretations have been highly speculative. Such is the lot of all factor analysts who use the tool in exploratory fashion. But I am not so much interested in what the most appropriate interpretations are as I am in the fact that substantial and interpretable differences exist. Now what if we had just proceeded along traditional lines and intercorrelated the 33 influence items across all teachers regardless of their school membership (the "total" analysis)? Table 13 presents the cluster analysis of the total correlation matrix, Figure 3

- -
Insert Table 13 & Figure 3 about here
- -

depicts the plot of eigenvalues against varying numbers of factor extractions, and Table 14 shows the results of the factor analyses. Comparing these

- - - - - - - - - - - - - - - -
Insert Table 14 about here
- - - - - - - - - - - - - - - -

data with the analogous results for the between and pooled within analyses, it is clear that the total analysis is much more similar to the pooled within analysis. Only 2-3% more variance is accounted for in the total analysis and the various factor patterns in Table 14 clearly imitate those in Table 12.

It is important to remember, however, that this is an empirical finding, not a mathematically derivable result. Certainly the magnitude of the total correlations can usually be expected to be somewhere intermediate to those of the pooled within (the smallest) and between (the largest) correlations. (See footnote 12.) But whether the total correlations will be more like the between or pooled within correlations depends upon the relative amount of between to within group variation. (See the discussion in the first two sections of this paper.) Although only a rough guide to the multivariate structures, the relatively small eta coefficients for most items (see Table 8) suggest the empirical finding, viz., that the total correlation matrix will be much more influenced by individuals (teachers) than by groups (schools).[15]

In spite of this general outcome, however, there still remain sufficient differences in clustering and factoring tendencies of the total and pooled within correlation matrices. For example, the 4-factor solutions given in Tables 14 and 12 indicate that the "staff meeting" items are more reluctant to separate from the "general school" cluster in the total analysis than in the pooled within analysis. This is a good example of how the total

analysis can be substantially influenced by a particularly strong between group tendency. On the other hand, the dichotomy between professional staff and teaching assistants in the "selection and evaluation" cluster shows up sooner in the total analysis. Thus it is clear that were I to settle on the 4-factor solution as the most reasonable interpretation of the dimensionality in the 33 teacher influence items, I would not be reporting substantively identical findings and would probably undergo a bit more consternation as an interpreter of the total analysis. As more factors are interpreted, the similarities between both analyses out number the differences. Nevertheless, the several differences are not trivial. In the 9-factor solution, for example, the "parent" cluster disappears from the total analysis, again as a direct result of the relatively low correlation between these two items using school as the unit of analysis. Instead a new cluster emerges in the total analysis (factor I) which, although interpretable ("student management"), is probably more an artifact of the number of factors rotated. (See item intercorrelations in Table 13.)

So we have a clear illustration of the statistical and substantive differences among the total-, between- and within-group approaches even when between-group differences are relatively small. Now what is the researcher's next step in dealing with these data? Nothing more or less than acknowledging the different measurement perspectives implied by the different approaches, selecting those most appropriate to the research study, and taking his/her best shot at developing scales suggested by both the factor/cluster analyses and the substantive intent of the items.[16] The scaling possibilities represented in Tables 15 & 16 are my best shots at the moment and pretty

- - - - - - - - - - - - - - - - - - -
Insert Tables 15 & 16 about here
- - - - - - - - - - - - - - - - - - -

much parallel the above discussion. For the pooled within analysis (Table 15), three major scaling possiblities are suggested: (a) 3-scales: W(123), W4 and W(56), (b) 4-scales: W(123), W4, W5 and W6 and (c) 6-scales: W1-W6. (The three 2-item scales below the dashed line are optional, regardless of the above possibilities.) Three major scaling possibilities are also suggested for the between analysis (Table 16): (a) 3 scales: B(123), B4 and B5, (b) 4 scales: B1, B(23), B4 and B5 and (c) 5 scales: B1-B5. (The order in which I have separated primary clusters into secondary clusters is based on the relative strengths of intra-cluster associations; this is mirrored quite closely by the order of item separations in the factor analyses as more factors are rotated.)

For the sake of illustration, suppose the 3-scale interpretations are selected, and we wish to look at internal co sistency coefficients and scale intercorrelations. Obviously for these computations we must use the between covariance matrix for the set of "between scales" and the pooled within covariance matrix for the set of "pooled within scales". Table 17 presents these data. Considering our approach to scale construction (see footnote 16), these results are not unexpected, viz., substantial alphas and moderate (but not unduly high) scale intercorrelations.[17] These are direct results from the psychometric phase of our study.

Also included in Table 17 are the intercorrelations between schools using the set of pooled within scales. It is important to reemphasize (refer back to Table 5) the potential to correlate pooled within scales both within and between schools. The former matrix is a product of the psychometric phase; but the latter matrix represents findings from the study phase. School means, based on pooled within scales, must be interpreted as averages of internally consistent individual perceptions within the school and treated

as such. Between scales, however, do not lend themselves to pooled within correlations--it makes no sense to score individuals on scales derived by ignoring individual variation. The between scales measure systemic constructs which are operationally defined by item means and should be treated as such. The intercorrelations of these scales are products of the psychometric phase.

Finally, in order not to further complicate matters, I have witheld discussion of differences due to other relevant factors in the study, e.g., level of schooling (elementary, junior high and senior high schools). In effect, we add a nesting factor to the design back in Table 3, "correcting" for both school and level effects in the pooled within analyses, pooling the between analysis, and adding a between levels analysis. This decision has a lot to do with the homogeneity of the within school correlation matrices--an assumption required for a sensible interpretation of the pooled within analysis. For example, an inspection of the cluster analyses of the 33 items within each school separately revealed primary clustering generally similar to the pooled within analysis.

As always, however, there were exceptions. And looking at pooled within analyses performed separately by level revealed several interesting results: (1) the 4-factor/cluster interpretation of the pooled within analysis over all schooling levels adequately describe each level taken separately and would seem to be viable scales for comparing schools across levels. (2) The senior high and junior high pooled within analyses are nearly identical to each other and therefore to the analysis over all schooling levels. (Of the total 1416 teachers in the study, 75% are from

secondary schools.) (3) The elementary school analysis, although showing very similar subclustering patterns, also shows a fundamentally different primary clustering pattern. Looking at the pooled within cluster analysis for elementary schools in Table 18, we see a more distinct separation between the issues of immediate classroom impact and all other items. For example, note how item 6 ("daily school schedule for students") now becomes part of the "curriculum, instruction and pupil behavior" cluster. In secondary schools, teachers always associated this item more with "general school" items. These results seem quite sensible given the departmentalized and period structure of secondary schools and the all-day, intact classroom structure of most of the elementary schools in our sample. Overall, one gets the feeling that perhaps elementary teachers (in contrast to secondary teachers) more easily compartmentalize their influence in the decision-making process as being either intra- or extra-classroom related. Again, this would seem reasonable given the more centralized operating process of elementary schools and more de-centralized operating process of secondary schools.

Aside from the validity of these speculations, we see once again that different partitionings of the total multivariate structure of a set of items do not produce the same result. Inevitably, the researcher must face the conceptual costs & benefits in utilizing these analyses <u>as descriptions</u> of group differences versus utilizing them for deriving constructs--constructs assumed to have consistent meaning for each group and on which groups can be compared. This is a methodological issue far more profound than the unit-of-analysis problem. (See footnote 9.)

Conclusion

It has been a long journey, but I hope my decision to thoroughly exemplify what could easily have remained abstract was worth the trip. We have seen (1) mathematical support for the potential between/within differences (2) an almost universal avoidance of these differences in one fairly large body of literature and (3) empirical evidence of these differences with respect to just one "climate-like" instrument.[18]

However, understanding these analytical differences and internalizing the conceptual implications with respect to a particular purpose and set of items are two processes which do not necessarily go hand in hand--at least not for this researcher. I am still struggling, for example, to internalize what it is that is being measured by "between scales", especially in the presence of substantial within group item variances. This struggle is hard to avoid when pressed with the inevitable question from my fellow researchers, "So what scales should I use--the between or within?" I follow this with the inevitable response, "Well what are you trying to measure?" The ensuing discussions seem to be easier when within analyses emerge as the most appropriate choice. For example, one colleague is currently investigating principal-staff congruencies on a number of climate-type measures and the relationship between congruency and a host of other school related variables. Should between or within scales be used to assess principal-staff congruency? As we pursued this question it became increasingly apparent that my colleague was interested in comparing the principal's phenomonological perspective with that of the staff's--in effect, one or more measures of "congruency" would be derived by evaluating the position of the principal's perception (scale score) in the <u>distribution</u> of teachers' perceptions (scale scores). All

of this leads to the necessity for a kind of substantive integrity for each person's score and thus to the pooled within analysis.

It is interesting to speculate about what our own humanness does to us as researchers interpreting data. We are much more apt to perceive and interpret phenomena like the humans in the system rather than like the system "perceives" and "interprets" phenomena--we feel like the people in a situation and not like the situation itself. But when researchers analyze system level effects, they must explain the emergent dimensions systemically, i.e., they must interpret and understand the item associations as though they were a context rather than members of the context. No wonder researchers might find it somewhat alien to work with between analyses. In any event, I challenge all of you out there to resist the temptation of total analyses and try your hand interpreting, when appropriate, between and within analyses. Perhaps if enough of us begin to take more seriously the consequences of our measurement procedures, we will not only bring operationalization more in line with conceptualization, but also remind ourselves and our fundors about the benefits to be accrued by increasing the ratio of time spent on the processes of inquiry to that spent on the outcomes.

Footnotes

1. As always, responsibility for this work remains solely with the author. But an acknowledgement is due the staff at IDEA, many of whom have helped me grapple with the issues raised in this paper. Special thanks are due Jerry Engstrom and Mitch Nides who served as sounding boards for many of the ideas expressed herein and who gathered together and helped synthesize the relevant climate literature. I also wish to thank Lee Cronbach, Tom Knapp, Leigh Burstein, and Noreen Webb for their valuable comments on this and earlier drafts of the manuscript.

2. The pooled within-group correlation matrix is actually obtained by subtracting the group mean on each item from every individual's score on that item within each group and then computing the correlations among items across all individuals using these residual scores.

3. Although not derived for the same purpose, a variant of this formula was derived much earlier by Wright (1917). See Knapp (1977) for a good summary of the literature, formulas and implications for bivariate correlations.

4. The formulas for ρ_b and η_X^2 are as follows:

$$\rho_b = \frac{\sum_j n_j(\mu_{Xj}-\mu_X)(\mu_{Yj}-\mu_Y)}{\sqrt{\sum_j n_j(\mu_{Xj}-\mu_X)^2 \sum_j n_j(\mu_{Yj}-\mu_Y)^2}}$$

$$\eta_X^2 = \frac{\sum_j n_j(\mu_{Xj}-\mu_X)^2}{\sum_i \sum_j (X_{ij}-\mu_X)^2}$$

where μ_X and μ_Y are the overall means of X and Y respectively.

5. This is a critical point since much of the argument to follow in this paper will call for an understanding of concepts and purpose in coping with unit-of-analysis issues. It will not be necessary that this understanding be in the form of a theory, i.e., "a set of interrelated constructs (concepts), definitions, and propositions that presents a systematic view of phenomena by specifying relations among variables, with the purpose of explaining and predicting the phenomena." (Kerlinger, 1973, p.9) It will only be necessary that enough conceptualization or pre-thinking has been done in order to cogently argue for the selection of one or more units for analysis--in effect, a reasonable point-of-view is required, not a full blown theory in the traditional, paradigmatic sense of the term.

6. An interesting complication arises depending upon one's conceptualization of the role played by persons within groups. A purely systemic view might regard the measurement operation in terms of those, and only those, people within the organization at the time of assessment -- this defines the P factor as fixed within groups. However, a perspective allowing for both psychological and organizational variance in the measurement operation would probably regard the people within the organization as but one of many possible samples -- this defines the P factor as random within groups. The former view results in a simplified $\hat{\alpha}_b$ formula which I suspect has been the one used by any researchers who have done between analyses. The latter view complicates matters since within group variation becomes part of the expected between group variation, and $\hat{\alpha}_b$ must be adjusted accordingly. Although these conceptual differences are critical, (1) they have no effect on the formula for $\hat{\alpha}_w$ and (2) the difference in computational results for fixed and random estimates of α_b are not great with the fixed result being slightly less than the random result.

7. More information on the background and current activities of A Study of Schooling can be found in the series of four articles to be published in the Phi Delta Kappan beginning with Goodlad, Sirotnik and Overman (1979).

8. A reasonably comprehensive review of the "climate" literature was conducted for the period including the years 1965 through 1977. The review included a computer search of the Eric and Psychological Abstracts files using key phrases such as work environment, organizational climate and educational environment. Additionally, the Current Index of Journals in Education and the Education Index were examined for relevant sources. Finally, the Academy of Management Journal, Administrative Science Quarterly, Review of Educational Research, and Review of Research in Education, were intensively canvassed over the above period. These efforts produced over 150 reports, published articles and/or books dealing with the study of organizational climate variables. Of these, over seventy studies pertained directly to the development of and/or psychometric analysis of climate inventories, discussions of the theoretical conceptualization of climate constructs, and studies of the correlates of measured climate constructs. These studies were reviewed in detail and form the basis for our literature review.

9. An interesting parallel emerges here with the "emic" and "etic" concepts first introduced in linguistics (Pike, 1967) and translated for use in cultural anthropological research. (See, for example, Goodenough, 1970.) Briefly the emic view suggests the study of human behavior within its own setting, while the etic view suggests the study of those aspects of human behavior which might be generalizeable, e.g., across cultural groups. Obviously traditional research methods and analytic techniques have grown out of etic mentality. Perhaps I might be stretching the concepts a bit, but the kind of dilemma being alluded to here might be thought of as an analytic counterpart to a more emic approach to data analysis. Although less difficult to ignore in the context of dealing with unit-of-analysis issues, this problem is much more fundamental to the research process per se.

10. In our review of the climate literature, we found only one instance where, during the study phase, something other than the arithmetic average was used for the aggregate statistic (Mansfield and Payne, 1977, used scale variances).

11. It should be noted that a statistical problem can occur when one analyzes between correlation matrices. It can be shown mathematically that the number of independent clusters of items can be no greater than the number of cases or items--whichever is smallest--entering into the computation of the correlations. This is not a problem for our between matrix based on all 35 schools. However, when the number of cases is smaller than the number of items (as for the separate between matrices for each schooling level), an artifactual impression of more interitem dependence is created, and factor analyses can often produce distorted factor structures. Yet, the most outstanding clusters can still be detected when the correlation matrices are inspected.

12. Aside from the linear dependency problem discussed in footnote 11, it can be shown mathematically that between correlations are always larger than the pooled within or total correlations whenever the pooled within is less than the total correlation. This is exactly what is expected when groups (e.g., schools) are formed on a non-random basis. In effect, homogeneous sub-groups will generally be less variable in their perceptions of phenomena specific to that subgroup. Combining subgroups into a single group will create the impression of more variability in perceptions, simply because of non-random differences between groups. Thus correlations computed across all individuals and groups will tend to be higher than those computed for individuals within any given group.

13. Both oblique and orthogonal solutions produced nearly the same results when rotating 3 to 9 factors for either type of correlation matrix.

14. This should not be interpreted as a general finding for between versus within analyses; it is only a finding for our 33 "teacher influence" items. A crude but descriptive index of how closely a simple structure is obtained in approximating the item dimensionality can be given as the proportion of ideally zero (non-Xs in Tables 11 and 12) loadings which are "near-zero" (say less than .25). For example, in the 4-factor solution, this proportion is .81 for the pooled within analysis and .63 for the between analysis indicating the analytical clarity of the former analysis.

15. Given the likely assumption of quite small between-group variance components relative to within-group components for most perceptual-type items, it is a good guess that most scales developed in the organizational climate literature approximate pooled within results.

16. This approach to scale construction, i.e., a rational combination of information based upon empirical item associations and a priori (intuitive/conceptual) constructs will necessarily result in correlated dimensions. We feel that striving for orthogonality is not only futile from the standpoint of replication, but also irrational in view of the fact that our conceptualization of the dimensionality for a particular space (e.g., perceived organizational climate) is never orthogonal--our brains do not think orthogonally. To the extent that orthogonality exists, it should be discovered, not imposed. Clearly, however, it is not critical to the main issues of this paper that one accept this approach to scale construction.

17. The higher alphas for "between scales" do not necessarily make them better choices than the "within scales". Once again, fundamentally different constructs are being measured conceptually and operationally. The relative sizes of the alphas are a direct result of the relationship noted previously in footnote 12.

18. Although more complicated and difficult to report given the large number of items, we have uncovered equally large and interpretible between- and within-group differences in our analysis of a 113-item class climate questionnaire for students and a 100-item school climate questionnaire for teachers.

References

Anderson, G.J. & Walberg, H.J. Learning environments. In H.J. Walberg (Ed.) Evaluating educational performance. Berkeley, Ca.: McCutchan, 1974.

Bentzen, M.M. Changing schools: The magic feather principle. New York: McGraw-Hill, 1974.

Burstein, L. The use of data from groups for inferences about individuals in educational research. Unpublished Ph.D. dissertation, Stanford University, 1975.

Cohen, M.C. & Nagel, E. An introduction to logic and scientific method. New York: Horcourt, Brace & World, Inc., 1934.

Cronbach, L.J. (with assistance of Deken, J.E. & Webb, N.). Research on classrooms and schools: Formulation of questions, design and analysis. Occasional Paper of the Stanford Evaluation Consortium, Stanford University, 1976.

Cronbach, L.J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Dixon, W.J. & Brown, M.B. (Eds.) Biomedical computer programs P-series. Berkeley, Ca.: University of California Press, 1977.

Goodenough, W.H. Description and comparison in cultural anthropology. Chicago: Aldine Publishing Co., 1970.

Goodlad, J.I., Sirotnik, K.A., & Overman, B.C. A study of schooling: an overview. Phi Delta Kappan, 1979, 61, 174-178.

Guion, R.M. A note on organizational climate. Organizational behavior and human performance, 1973, 9, 120-125.

Halpin, A.W. & Croft, D.B. The organizational climate of schools. Chicago, Ill.: Midwest Administration Center, The University of Chicago, 1963.

Hartigan, J.A. Clustering Algorithms. New York: John Wiley & Sons, 1975.

Hayes, A.E. A reappraisal of the Halpin-Croft model of the organizational climate of schools. Paper presented at the American Educational Research Association Annual Meeting, 1973.

House, R.J. & Rizzo, J.R. Toward the measurement of organizational practices: Scale development and validation. Journal of Applied Psychology, 1972, 56, 388-396.

Howe, J.G. & Gavin, J.F. Organizational climate: A review and delineation. Technical Report No. 7402. Fort Collins, Co.: Colorado State University, Industrial Psychological Association of Colorado, 1974.

Jones, A.D. & James, L.R. Psychological and organizational climate: Dimensions and relationships. Technical Report No. 76-4. Fort Worth, Tex.: Texas Christian University, Institute of Behavior Research, 1976.

Kenney, J.B. & Rentz, R.R. The organizational climate of schools in five urban areas. The Elementary School Journal, 1970, 71, 61-69.

Kerlinger, F.N. Foundations of behavioral research. New York: Holt, Rinehart & Winston, 1973.

Kimpston, R.D. & Sonnabend, L.C. Public secondary schools. Urban Education, 1975, 10, 27-45.

Knapp, T.R. The unit-of-analysis problem in applications of simple correlation analysis to educational research. Journal of Educational Statistics, 1977, 2, 171-186.

Likert, J.G. & Likert, R. The Likert profile of a school. Ann Arbor, Mich.: Rensis Likert Associates, 1972.

Litwin, G. & Stringer, R. Motivation and organizational climate. Cambridge Mass.: Harvard University Press, 1968.

Mansfield, R. & Payne, R.L. Correlates of variance in perceptions of organizational climate. In D.S. Pugh & R.L. Payne (Eds.) Organizational behavior in its context, Westmead, England: Saxon House, Teakfield Limited, 1977.

Marks, E. Some results relating to the choice of an experimental unit in multivariate research. Multivariate Behavioral Research, 1970, 5, 507-520.

Mitchell, J.V. Dimensionality and differences in the environmental press of high schools. American Educational Research Journal, 1968, 5, 513-530.

Moos, R.H. Assessment of the psychosocial environments of community-oriented psychiatric treatment programs. Abnormal Psychology, 1972, 79, 9-18.

Moos, R.H. & Houts, P.S. Assessment of the social atmospheres of psychiatric wards. Journal of Abnormal Psychology, 1968, 73, 595-604.

Newman, J.E. Understanding the organizational structure--Job attitude relationship through perceptions of the work environment. Organizational Behavior and Human Performance, 1975, 14, 371-397.

Norman, W.T. On estimating psychological relationships: Social desirability and self-report. Psychological Bulletin, 1967, 67, 273-293.

Pace, C.R. CUES: College and university environment scales (technical manual). Princeton, New Jersey: Educational Testing Service, 1963.

Pace, C.R. & Stern, G.G. An approach to the measurement of psychological characteristics of college environments. Journal of Educational Psychology, 1958, 49, 269-277.

Payne, R.L., Fineman, S. & Wall, T.D. Organizational climate and job satisfaction: a conceptual synthesis. Organizational Behavior and Human Performance, 1976, 16, 45-62.

Payne, R.L. & Pheysey, D.C. G.G. Stern's organizational climate index: A reconceptualization and application to business organizations. Organizational Behavior and Human Performance, 1971, 6, 77-98.

Pike, K.L. Language in relation to a unified theory of the structure of human behavior. The Hague: Mouton, 1967

Robinson, W.S. Ecological correlations and the behavior of individuals. American Sociological Review, 1950, 15, 351-357.

Schneider, B. & Bartlett, C.J. Individual differences and organizational climate: I. The research plan and questionnaire development. Personnel Psychology, 1968, 21, 323-333.

Sirotnik, K.A., Nides, M.A. and Engstrom, G.A. Some methodological issues in developing measures of classroom learning environment. Studies in Educational Evaluation, 1980, 6, (3), in press.

Sockloff, A.L. Behavior of the product-moment correlation coefficient when two heterogeneous subgroups are pooled. Educational and Psychological Measurement, 1975, 35, 267-276.

Stern, G.G. People in context. New York: Wiley, 1970.

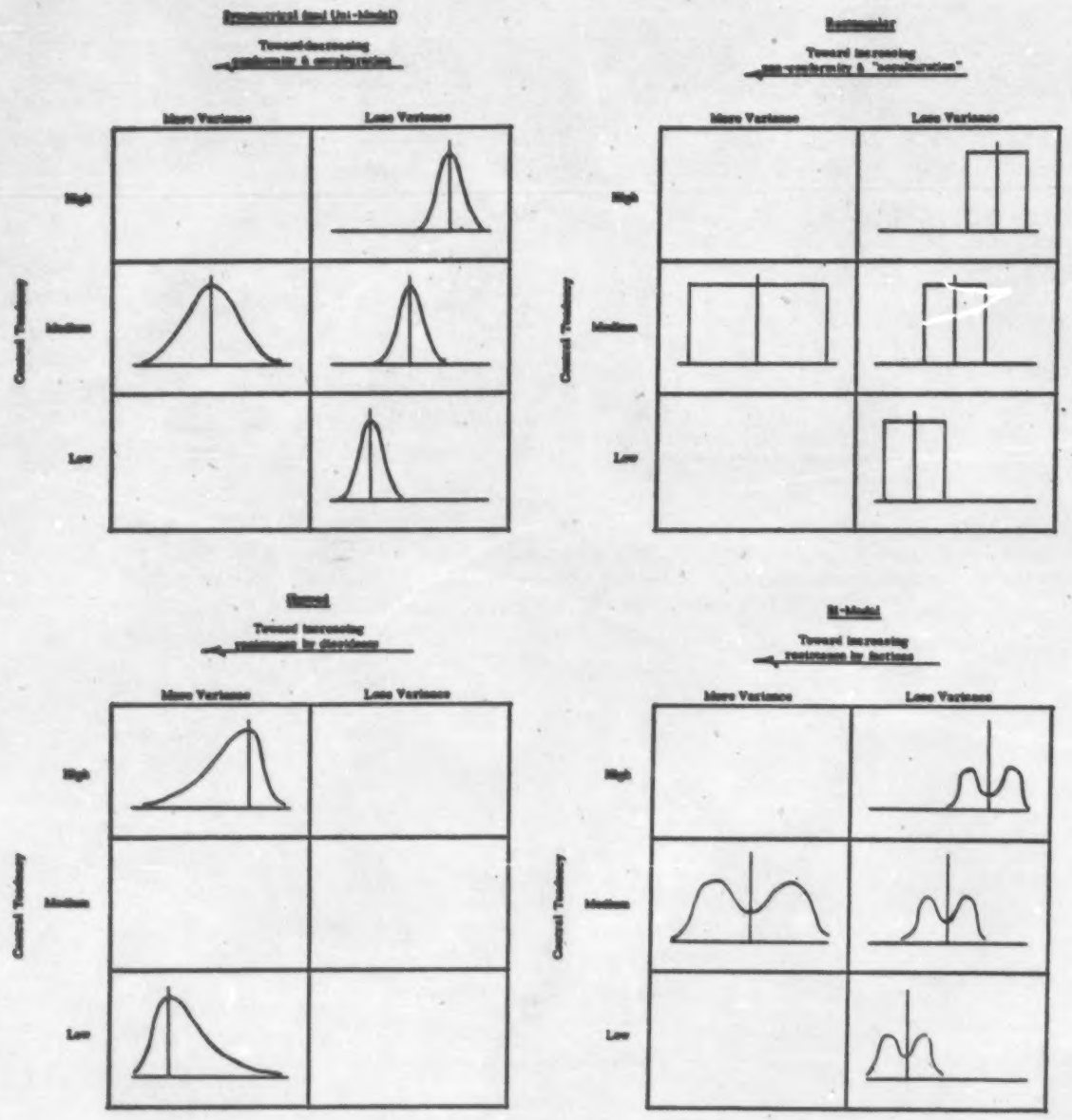
Taylor, J.C. & Bowers, D.G. Survey of Organizations. Ann Arbor: Center for Research on Utilization of Scientific Knowledge and the Institute for Social Research, 1972.

Walberg, J.J., Sorenson, J. & Fischbach, T. Ecological correlates of ambience in the learning environment. American Educational Research Journal, 1972, 9, 139-148.

Wright, S. The average correlation within subgroups of a population. Journal of the Washington Academy of Science, 1917, 7, 532-535.

Figure 1

The Distribution of Perceptions on a Climate Construct:
Idealized Distributions in Terms of Shape,
Variability and Central Tendency and Hypothetical Descriptions
(Missing cells due to limited ranges of typical Likert scales)



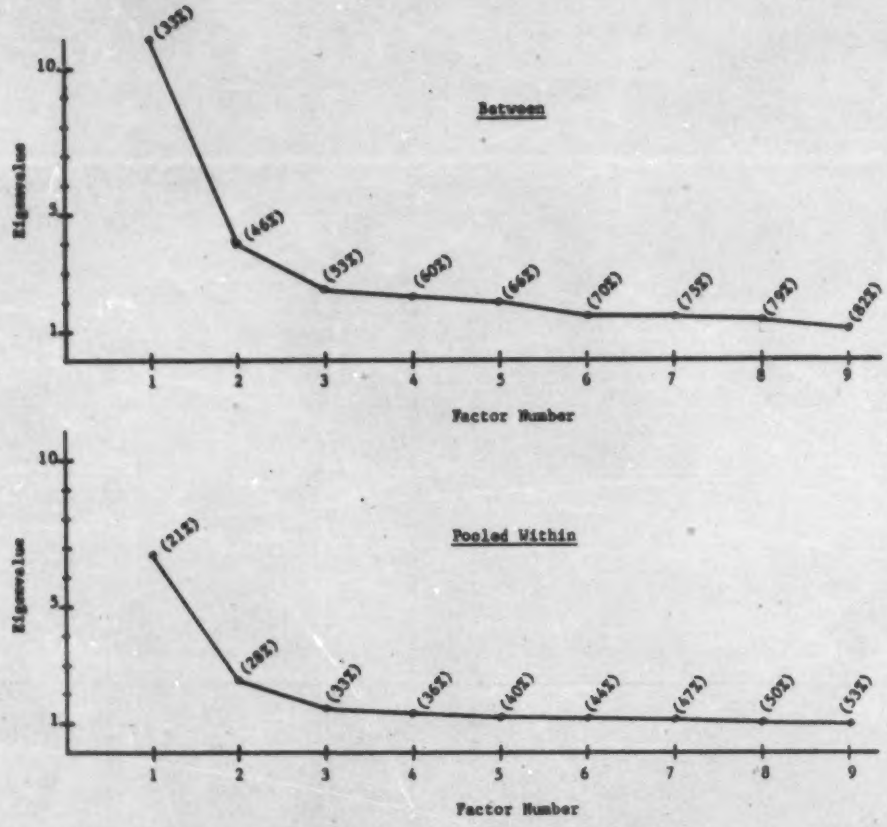


Figure 2. Plot of eigenvalues for varying numbers of unrotated principal components from between and pooled within correlation matrices for teacher influence items (cumulative % of variance accounted for indicated in parentheses).

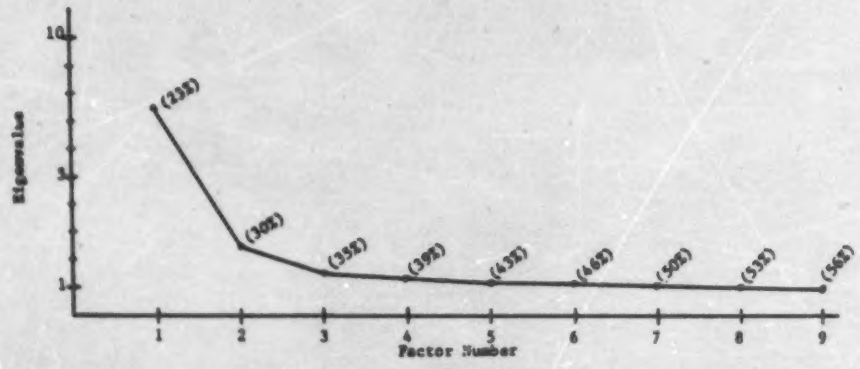


Figure 3. Plot of eigenvalues for varying numbers of unrotated principal components from the total correlation matrix for teacher influence items (cumulative % of variance accounted for indicated in parentheses).

Table 1a

Substantial Between Group Correlation
and Within Group Homogeneity

Group	Individual	Item Scores X	Item Scores Y	Within Group Correlation
1	1	1	1	.71
	2	2	1	
	3	2	2	
	4	2	3	
	5	3	3	
		$\mu_{X_1} = 2$	$\mu_{Y_1} = 2$	
2	1	2	3	.71
	2	2	4	
	3	3	4	
	4	4	4	
	5	4	5	
		$\mu_{X_2} = 3$	$\mu_{Y_2} = 4$	
3	1	3	3	.71
	2	3	4	
	3	4	4	
	4	5	4	
	5	5	5	
		$\mu_{X_3} = 4$	$\mu_{Y_3} = 4$	

$\rho_b = .87$ $\quad \eta_X^2 = .50$

$\rho_w = .67$ $\quad \eta_Y^2 = .62$

$\rho_t = .77$

Table 1b

Substantial Between Group Correlation
and Within Group Heterogeneity

Group	Individual	Item Scores X	Item Scores Y	Within Group Correlation
1	1	1	1	.71
	2	2	1	
	3	2	2	
	4	2	3	
	5	3	3	
		$\mu_{X_1} = 2$	$\mu_{Y_1} = 2$	
2	1	2	5	-.71
	2	2	4	
	3	3	4	
	4	4	4	
	5	4	3	
		$\mu_{X_2} = 3$	$\mu_{Y_2} = 4$	
3	1	3	5	0
	2	3	3	
	3	4	4	
	4	5	4	
	5	5	4	
		$\mu_{X_3} = 4$	$\mu_{Y_3} = 4$	

$\rho_b = .87$ $\quad \eta_X^2 = .50$

$\rho_w = 0$ $\quad \eta_Y^2 = .62$

$\rho_t = .48$

Table 1c

Zero Between Group Correlation
and Within Group Homogeneity

Group	Individual	Item Scores X	Item Scores Y	Within Group Correlation
1	1	1	3	.71
	2	2	3	
	3	2	4	
	4	2	5	
	5	3	5	
		$\mu_{X_1} = 2$	$\mu_{Y_1} = 4$	
2	1	2	1	.71
	2	2	2	
	3	3	2	
	4	4	2	
	5	4	3	
		$\mu_{X_2} = 3$	$\mu_{Y_2} = 2$	
3	1	3	3	.71
	2	3	4	
	3	4	4	
	4	5	4	
	5	5	5	
		$\mu_{X_3} = 4$	$\mu_{Y_3} = 4$	

$\rho_b = 0$ $\quad \eta_X^2 = .50$

$\rho_w = .67$ $\quad \eta_Y^2 = .62$

$\rho_t = .29$

Table 1d

Zero Between Group Correlation
and Within Group Heterogeneity

Group	Individual	Item Scores X	Item Scores Y	Within Group Correlation
1	1	1	3	.71
	2	2	3	
	3	2	4	
	4	2	5	
	5	3	5	
		$\mu_{X_1} = 2$	$\mu_{Y_1} = 4$	
2	1	2	3	-.71
	2	2	2	
	3	3	2	
	4	4	2	
	5	4	1	
		$\mu_{X_2} = 3$	$\mu_{Y_2} = 2$	
3	1	3	3	0
	2	3	5	
	3	4	4	
	4	5	4	
	5	5	4	
		$\mu_{X_3} = 4$	$\mu_{Y_3} = 4$	

$\rho_b = 0$ $\quad \eta_X^2 = .50$

$\rho_w = 0$ $\quad \eta_Y^2 = .62$

$\rho_t = 0$

Table 2

Total Analysis: n Persons Crossed With i Items Assuming No Grouping Effects Present

Source	df	Expected Mean Square
I	i-1	$\sigma^2_{PI} + n\sigma^2_I$
P	n-1	$\sigma^2_{PI} + i\sigma^2_P$
PI	(n-1)(i-1)	σ^2_{PI}

$$\hat{\alpha} = \frac{\hat{\sigma}^2_P}{\sigma^2_P + \frac{1}{i}\hat{\sigma}^2_{PI}} = \frac{MS_P - MS_{PI}}{MS_P}$$

Table 3

Between and Pooled Within
Analysis: p Persons Nested Within
g Groups Crossed With i Items

Source	df	Expected Mean Square	
		P Fixed	P Random
I	i-1	$p\sigma^2_{GI} + gp\sigma^2_I$	$\sigma^2_{PI} + p\sigma^2_{GI} + gp\sigma^2_I$
P	g(p-1)	$\sigma^2_{PI} + i\sigma^2_P$	$\sigma^2_{PI} + i\sigma^2_P$
PI	g(p-1)(i-1)	σ^2_{PI}	σ^2_{PI}
G	g-1	$p\sigma^2_{GI} + pi\sigma^2_G$	$\sigma^2_{PI} + p\sigma^2_{GI} + i\sigma^2_P + pi\sigma^2_G$
GI	(g-1)(i-1)	$p\sigma^2_{GI}$	$\sigma^2_{PI} + p\sigma^2_{GI}$

Alpha Coefficient	Variance Component Definition	Estimation Formula	
		P Fixed	P Random
α_w	$\dfrac{\sigma^2_P}{\sigma^2_P + \frac{1}{i}\sigma^2_{PI}}$	$\dfrac{MS_P - MS_{PI}}{MS_P}$	same
α_b	$\dfrac{\sigma^2_G}{\sigma^2_G + \frac{1}{i}\sigma^2_{GI}}$	$\dfrac{MS_G - MS_{GI}}{MS_G}$	$\dfrac{(MS_G - MS_{GI}) - (MS_P - MS_{PI})}{MS_G - MS_P}$

Table 4

A Total Analysis (gp Persons Crossed With i Items) When Grouping Effects Are Really Present

When P Is Fixed

Source	df	Expected Mean Square
I	i-1	$p\sigma^2_{GI} + gp\sigma^2_I$
P	gp-1	$\frac{g(p-1)}{gp-1}(\sigma^2_{PI} + i\sigma^2_P) + \frac{p(g-1)}{gp-1}(\sigma^2_{GI} + i\sigma^2_G)$
PI	(gp-1)(i-1)	$\frac{g(p-1)}{gp-1}\sigma^2_{PI} + \frac{p(g-1)}{gp-1}\sigma^2_{GI}$

$$\hat{\alpha}_t = \frac{MS_P - MS_{PI}}{MS_P} = \frac{[g(p-1)\hat{\sigma}^2_P + p(g-1)\hat{\sigma}^2_G]}{[g(p-1)\hat{\sigma}^2_P + p(g-1)\hat{\sigma}^2_G] + \frac{1}{i}[g(p-1)\hat{\sigma}^2_{PI} + p(g-1)\hat{\sigma}^2_{GI}]}$$

When P Is Random

Source	df	Expected Mean Square
I	i-1	$p\sigma^2_{GI} + gp\sigma^2_I$
P	gp-1	$(\sigma^2_{PI} + i\sigma^2_P) + \frac{p(g-1)}{gp-1}(\sigma^2_{GI} + i\sigma^2_G)$
PI	(gp-1)(i-1)	$\sigma^2_{PI} + \frac{p(g-1)}{gp-1}\sigma^2_{GI}$

$$\hat{\alpha}_t = \frac{MS_P - MS_{PI}}{MS_P} = \frac{\hat{\sigma}^2_P + \frac{p(g-1)}{gp-1}\hat{\sigma}^2_G}{[\hat{\sigma}^2_P + \frac{p(g-1)}{gp-1}\hat{\sigma}^2_G] + \frac{1}{i}[\hat{\sigma}^2_{PI} + \frac{p(g-1)}{gp-1}\hat{\sigma}^2_{GI}]}$$

Table 5

Conceivable Statistical Aggregates During the Study Phase for Scales Emerging from the Psychometric Phase for Three Possible Units of Analysis

	Study Phase		
Psychometric Phase	Individuals Across Groups	Individuals Within Groups	Groups
Individuals across groups	Individual composite scores	Individual composite scores	Any parameter of the score distribution
Individuals within groups	?	Individual composite scores	Any parameter of the score distribution
Groups	?	?	Only a composite of the statistical aggregate used at the item level

? = Analyses fitting these cells during the study phase are inconceivable

Table 6

Representative Selection of Instruments and Psychometric Methods Originating in Industrial and Educational Organization Settings

Setting, Study & Instrument	Within (a single organization)	Total (Across individuals across organizations)	Between (Item or scale means)	Within (Pooled over 2 or more organizations)
Industrial:				
House and Rizzo (1972) Organization Description Questionnaire	X			
Litwin & Stringer (1968) Form B (unnamed)		X		
Moos & Houts (1965) Ward Atmosphere Scale		X		
Newman (1975) Perceived Work Environment	X			
Payne & Pheysey (1971) Business Organization Climate Index		X		
Schneider & Bartlett (1968) Agency Climate Questionnaire		X		
Taylor & Bowers (1972) Survey of Organizations			X	
Educational:				
Halpin & Croft (1963) Organizational Climate Description Questionnaire[a]		X		
Kimpston & Sonnabend (1975) Organizational Health Description Questionnaire		X		
Likert & Likert (1972) Likert School Profiles				
Pace (1963) College & University Environment Scales[b]			X	
Stern (1970) Organizational Climate Index	X	X		

a Many other studies have investigated this instrument invariably using the total correlation matrix. (See, for example, Kenney & Rantz, 1970 and Hayes, 1973.)

b This instrument (adapted from the College Characteristics Index, Pace and Stern, 1958) is designed to assess students' perceptions of their learning environment, technically placing it in a different class from those instruments designed to assess employees' perceptions of their work environment. The unit-of-analysis issues are identical, however, and the CUES instrument is included here because of its unique psychometric treatment.

Table 7

"Teacher Influence"
Items as Used in the Study
of Schooling

The responsibilities that teachers have vary from school to school. Sometimes these responsibilities are small in number, sometimes they are large in number. Below is a list of some of the things about which teachers may help make decisions. Please indicate how much influence the teachers at your school have in decisions made about each of the following:

	A lot of influence	Some influence	No influence
1 Changes in curriculum	O	O	O
2 Instructional methods that are used in classrooms	O	O	O
3 Standards of pupil behavior in their own classrooms	O	O	O
4 Standards of pupil behavior in halls and on playground	O	O	O
5 Daily schedule in their own classroom	O	O	O
6 Daily school schedule for students	O	O	O
7 Special behavior problems with individual pupils	O	O	O
8 Special all school affairs, such as open house, assemblies, etc.	O	O	O
9 Committing the staff to participate in special projects or innovations	O	O	O
10 Community relations policy	O	O	O
11 School publications	O	O	O
12 Unusual problems that affect the whole school	O	O	O
13 Time of staff meetings	O	O	O
14 Content of staff meetings	O	O	O
15 The way in which staff meetings are conducted	O	O	O
16 Arrangements for parent conferences	O	O	O
17 Assignments for teacher duties outside of classrooms (yard duty, etc.)	O	O	O
18 Planning social gatherings of school staff	O	O	O
19 Standards of dress for pupils	O	O	O
20 Standards of dress for staff	O	O	O
21 Assigning pupils to classes	O	O	O
22 Assigning teachers to classes	O	O	O
23 Ways of reporting pupil progress to parents	O	O	O
24 Preparing the school budget	O	O	O
25 Managing the funds available for instructional purposes	O	O	O
26 Selecting volunteer teaching assistants	O	O	O
27 Selecting paid teaching assistants	O	O	O
28 Selecting part-time teachers for the school staff	O	O	O
29 Selecting full-time teachers for the school staff	O	O	O
30 Evaluating the performance of teaching assistants	O	O	O
31 Evaluating the performance of full-time teachers	O	O	O
32 The dismissal and/or transfer of teachers	O	O	O
33 Selecting administrative personnel to be assigned to the school	O	O	O

Table 8

Estimated Variance Components
Between (σ_B^2) and Within (σ_W^2) Schools
and Correlation Ratios (η^2)
for each of the 33 Teacher Influence Items

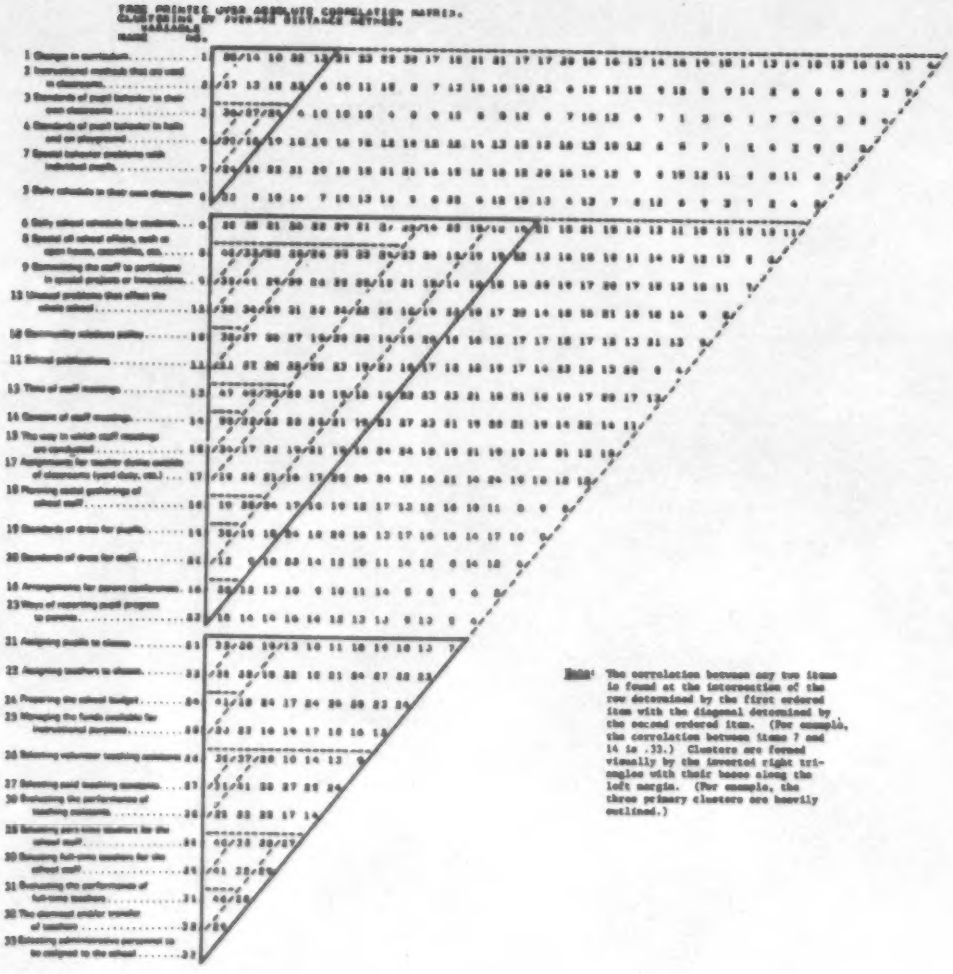
Item	Estimates			Item	Estimates		
	$\hat{\sigma}_B^2$	$\hat{\sigma}_W^2$	$\hat{\eta}^2$		$\hat{\sigma}_B^2$	$\hat{\sigma}_W^2$	$\hat{\eta}^2$
1	.0779	.3131	.198	18	.1138	.3886	.225
2	.0176	.2333	.070	19	.0446	.3113	.124
3	.0138	.2147	.060	20	.0558	.6039	.084
4	.0937	.2932	.241	21	.0765	.2997	.202
5	.0300	.3680	.075	22	.0402	.3014	.118
6	.0831	.3749	.181	23	.0438	.4417	.090
7	.0439	.2801	.134	24	.0484	.2080	.188
8	.0796	.3563	.181	25	.0565	.3246	.147
9	.0483	.3334	.126	26	.0822	.3852	.176
10	.0446	.3024	.128	27	.0184	.1652	.100
11	.0539	.3679	.128	28	.0135	.1189	.101
12	.0467	.2792	.143	29	.0432	.1606	.211
13	.0416	.3115	.117	30	.0398	.3808	.095
14	.0552	.2832	.163	31	.0097	.1528	.059
15	.0415	.2759	.130	32	.0068	.1146	.055
16	.0624	.3705	.143	33	.0015	.0510	.029
17	.1002	.2991	.250				

Table 9

Cluster Analysis of the Between Correlation Matrix

Key: The correlation between any two items is found at the intersection of the row determined by the first ordered item with the diagonal determined by the second ordered item. (For example, the correlation between items 7 and 14 is .33.) Clusters are formed visually by the inverted right triangles with their bases along the left margin. (For example, the three primary clusters are heavily outlined.)

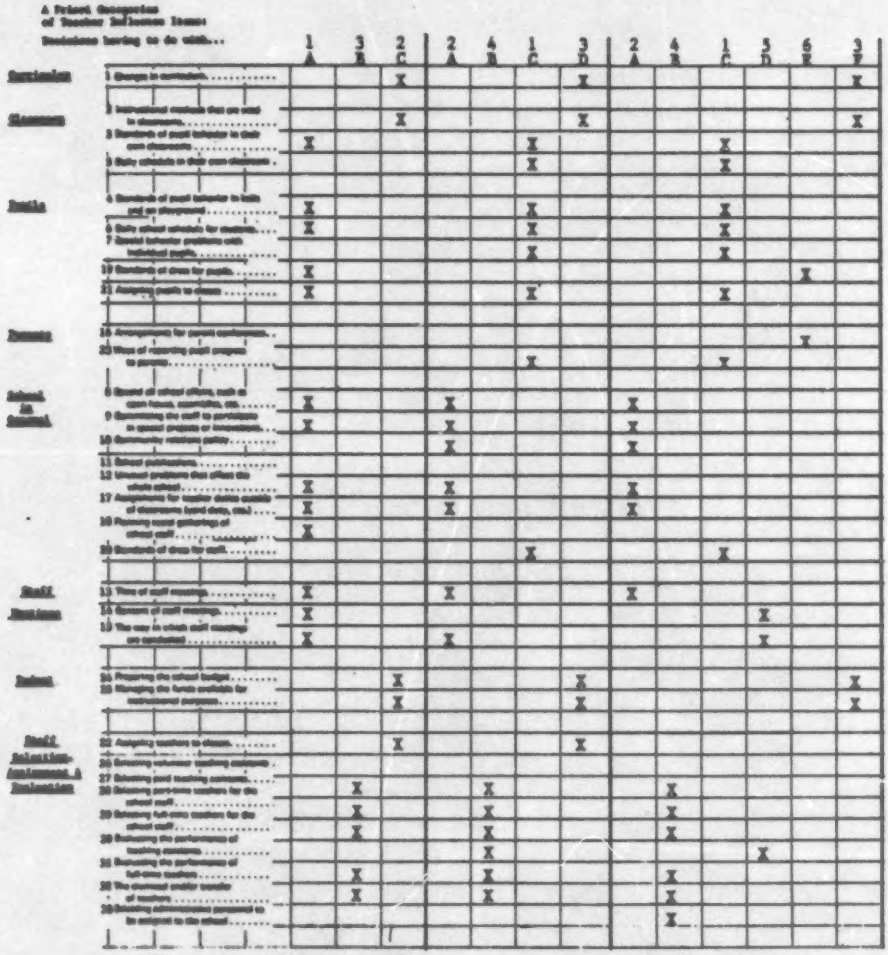
Table 10
Cluster Analysis of the Pooled Within Correlation Matrix



Note: The correlation between any two items is found at the intersection of the row determined by the first ordered item with the diagonal determined by the second ordered item. (For example, the correlation between items 7 and 14 is .33.) Clusters are formed visually by the inverted right triangles with their bases along the left margin. (For example, the three primary clusters are heavily outlined.)

Table 11

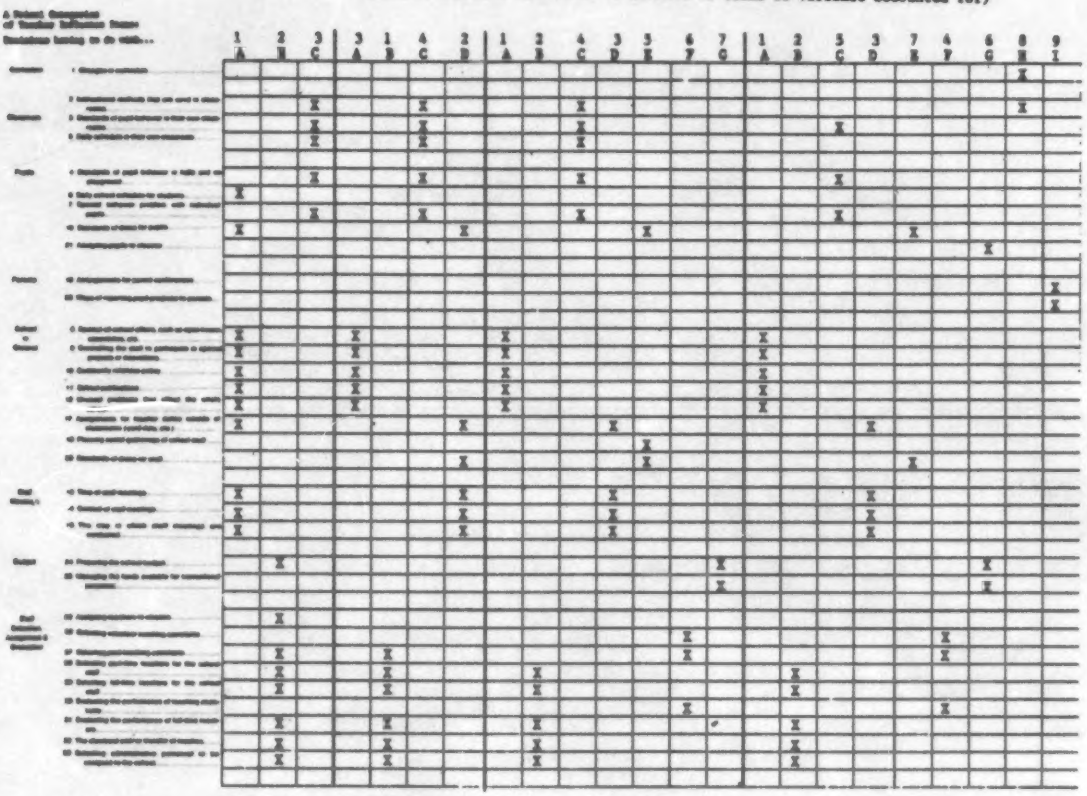
3-,4- and 6-Factor Solutions for the Between Correlation Matrix
(Numbers indicate rank order of factors in terms of variance accounted for)


A Priori Categories of Teacher Influence Items:
Decisions having to do with...
1 A
3 B
2 C
2 A
4 B
1 C
3 D
2 A
4 B
1 C
5 D
6 E
3 F
Curriculum
1 Changes in curriculum
Classroom
3 Instructional methods that are used in classrooms
2 Standards of pupil behavior in their own classrooms
3 Daily schedule in their own classrooms
Pupils
4 Standards of pupil behavior in halls and on playground
6 Daily school schedule for students
7 Special behavior problems with individual pupils
10 Standards of dress for pupils
11 Assigning pupils to classes
Parents
12 Arrangements for parent conferences
13 Type of reporting pupil progress to parents
School in General
8 Speed of school affairs, such as open house, assemblies, etc.
9 Encouraging the staff to participate in special projects or improvements
10 Community relations policy
11 School publications
12 Unusual problems that affect the whole school
17 Assessment for teacher duties (grades of classrooms, typed data, etc.)
18 Planning social gatherings of school staff
19 Standards of dress for staff
Staff Meetings
15 Time of staff meetings
14 Amount of staff meetings
15 The way in which staff meetings are conducted
Budget
16 Preparing the school budget
18 Managing the funds available for instructional purposes
Staff Selection-Assignment & Evaluation
22 Assigning teachers to classes
26 Securing substitute teaching assistants
27 Securing paid teaching assistants
28 Selecting part-time teachers for the school staff
29 Selecting full-time teachers for the school staff
30 Evaluating the performance of teaching assistants
31 Evaluating the performance of full-time teachers
32 The standard and/or transfer of teachers
33 Selecting administrators proposed to be assigned to the school

Note: X = loadings > .55

Table 12

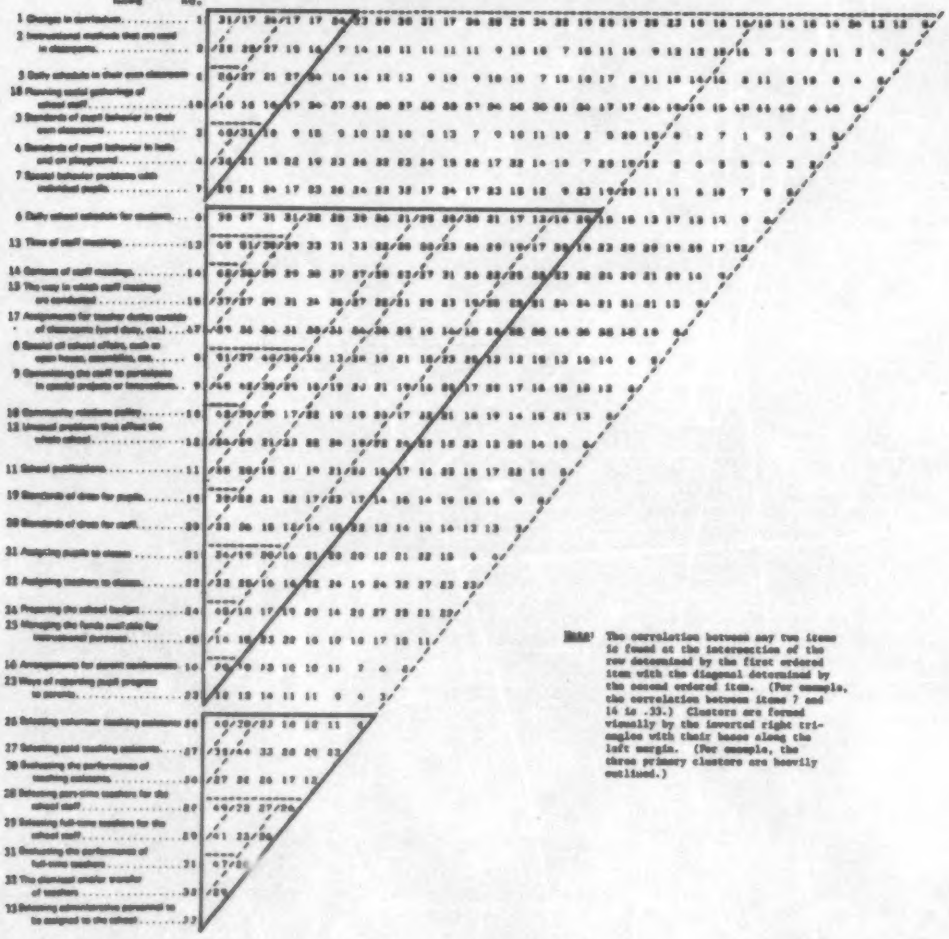
3-,4-, 7- and 9-Factor Solutions for the Pooled Within Correlation Matrix
(Numbers indicate rank order of factors in terms of variance accounted for)



Note: X = loadings ≥ .45

Table 13

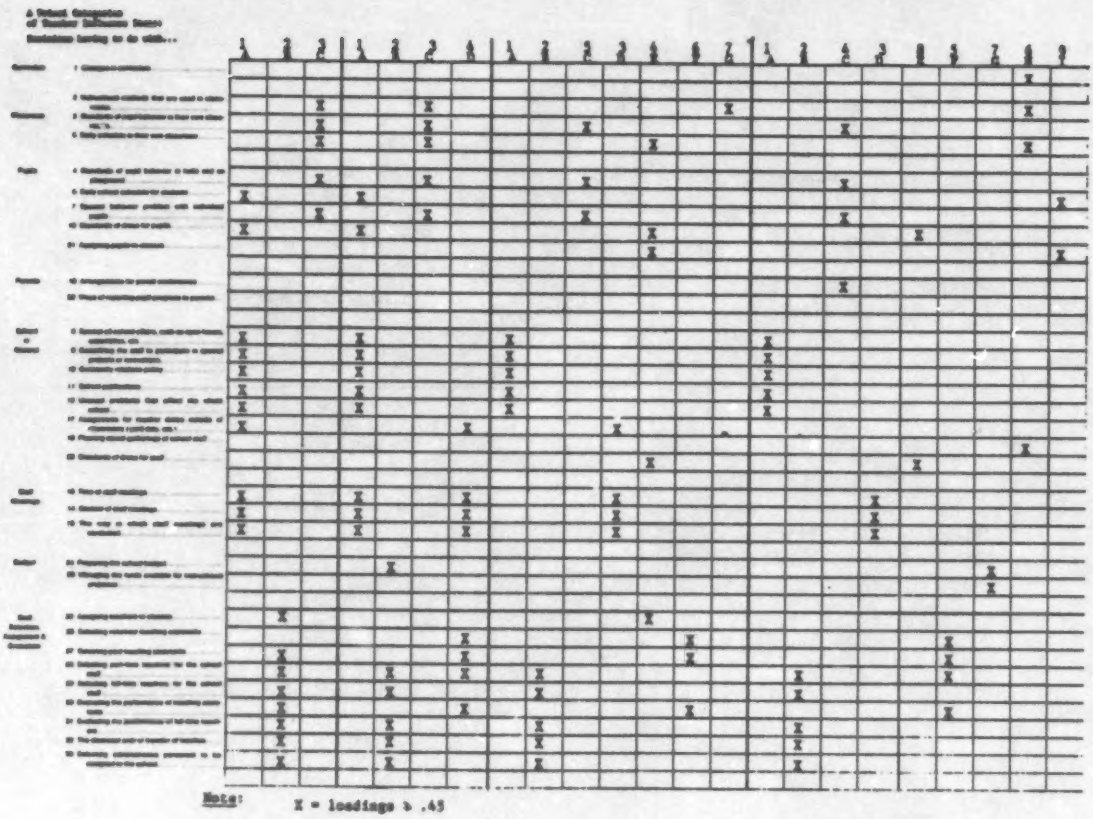
Cluster Analysis of the Total Correlation Matrix



Note: The correlation between any two items is found at the intersection of the row determined by the first ordered item with the diagonal determined by the second ordered item. (For example, the correlation between items 7 and 14 is .33.) Clusters are formed visually by the inverted right triangles with their bases along the left margin. (For example, the three primary clusters are heavily outlined.)

Table 14

3-, 4-, 7- and 9-Factor Solutions for the Total Correlation Matrix
(Numbers indicate rank order of factors in terms of variance accounted for)



Note: X = loadings ≥ .45

Table 15

Hierarchical Clustering & Scaling
Possibilities for Pooled Within Correlation Matrix

W(123). Selection, Evaluation & Fiscal Management

W1. Selection & Evaluation of Professional Staff

28 Selecting part-time teachers for the school staff.....
29 Selecting full-time teachers for the school staff.....
31 Evaluating the performance of full-time teachers.....
32 The dismissal and/or transfer of teachers.....
33 Selecting administrative personnel to be assigned to the school.....

W2. Selection & Evaluation of Teaching Assistants

26 Selecting volunteer teaching assistants
27 Selecting paid teaching assistants....
30 Evaluating the performance of teaching assistants.....

W3. Fiscal Management

24 Preparing the school budget.....
25 Managing the funds available for instructional purposes.....

W4. Curriculum, Instruction & Pupil Behavior

1 Changes in curriculum.....
2 Instructional methods that are used in classrooms.....
3 Standards of pupil behavior in their own classrooms.....
4 Standards of pupil behavior in halls and on playground.....
5 Daily schedule in their own classroom .
7 Special behavior problems with individual pupils.....

W(56). School Focused Issues

W5. Extra-Curricular & Community Related Issues & Activities

8 Special all-school affairs, such as open house, assemblies, etc.....
9 Committing the staff to participate in special projects or innovations...
10 Community relations policy.....
11 School publications.....
12 Unusual problems that affect the whole school.....

W6. Procedures & Content of Staff Meetings

13 Time of staff meetings.....
14 Content of staff meetings.....
15 The way in which staff meetings are conducted.....

Additional 2-Item Clusters:

Communication With Parents

16 Arrangements for parent conferences.
23 Ways of reporting pupil progress to parents.....

Pupil & Staff Class Assignments

21 Assigning pupils to classes.....
22 Assigning teachers to classes.....

Pupil & Staff Dress Code

19 Standards of dress for pupils.....
20 Standards of dress for staff.....

Table 16

Hierarchical Clustering & Scaling
Possibilities for Between Correlation Matrix

B(123). Pupil/Class-Related & School-General Issues

B1. Pupil Focused Issues

3 Standards of pupil behavior in their own classrooms
4 Standards of pupil behavior in halls and on playground
5 Daily schedule in their own classroom
6 Daily school schedule for students
7 Special behavior problems with individual pupils
16 Arrangements for parent conferences
19 Standards of dress for pupils
21 Assigning pupils to classes
23 Ways of reporting pupil progress to parents

B(23). School Focused Issues

B2. Extra-classroom Issues, Activities & Commitments

8 Special all-school affairs, such as open houses, assemblies, etc.
9 Committing the staff to participate in special projects or innovations
10 Community relations policy
12 Unusual problems that affect the whole school
13 Time of staff meetings
17 Assignments for teacher duties outside of classrooms (yard duty, etc.)
18 Planning social gatherings of school staff

B3. Procedures & Content of Staff Meetings

14 Content of staff meetings
15 The way in which staff meetings are conducted

B4. Resources & Curriculum

1 Changes in curriculum
2 Instructional methods that are used in classrooms
24 Preparing the school budget
25 Managing the funds available for instructional purposes

B5. Selection & Evaluation of Staff

27 Selecting paid teaching assistants
28 Selecting part-time teachers for the school staff
29 Selecting full-time teachers for the school staff
30 Evaluating the performance of teaching assistants
31 Evaluating the performance of full-time teachers
32 The dismissal and/or transfer of teachers

Table 17

Scale Intercorrelations and
Alpha Coefficients for Between and
Pooled Within Dimensions

Between Scales; Between Correlations

	B(123)	B4	B5	Alpha
B(123)	1.00	.46	.27	.92
B4		1.00	.27	.81
B5			1.00	.79

Pooled Within Scales; Pooled Within Correlations

	W(123)	W4	W(56)	Alpha
W(123)	1.00	.22	.46	.74
W4		1.00	.41	.62
W(56)			1.00	.80

Pooled Within Scales; Between Correlations

	W(123)	W4	(W56)
W(123)	1.00	.42	.53
W4		1.00	.59
W(56)			1.00

Table 18

Cluster Analysis of Pooled Within Correlation Matrix for Elementary Teachers Only